Hilbert transforms and maximal operators along planar vector fields

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Shaoming Guo

aus

Zhaoyuan, Shandong, China

Bonn, 2015

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms Universität Bonn

1. Gutachter: Prof. Dr. Christoph Thiele

2. Gutachter: Prof. Dr. Herbert Koch

Tag der Promotion: 6. Jul. 2015

Erscheinungsjahr: 2015

Abstract

In harmonic analysis, there is a conjecture (attributed to Zygmund in [21]) stating that the directional maximal operator along a Lipschitz planar vector field (defined as in (1.0.6)) is weakly bounded on $L^2(\mathbb{R}^2)$. In this thesis, we present some recent progress towards this conjecture and its singular integral variant, which is that the directional Hilbert transform along a Lipschitz vector field (defined as in (1.0.8)) is weakly bounded on $L^2(\mathbb{R}^2)$.

In Chapter 1 we will first state these two conjectures and explain some partial progress that has been made. Afterwards we will state the main results of the present thesis.

In Chapter 2 we will prove the L^2 boundedness of the directional Hilbert transform along planar measurable vector fields which are constant along suitable Lipschitz curves. Jones' beta numbers will play a crucial role when handling vector fields of the critical Lipschitz regularity.

In Chapter 3 we will generalise the L^2 bounds in Chapter 2 to L^p for all $p > 3/2$. To achieve this, we need to study a new paraproduct, which is indeed a one-parameter family of paraproducts, with each paraproduct living on one Lipschitz level curve of the vector field.

In Chapter 4, by using the techniques presented in Chapter 2 and 3, we will provide a geometric proof of Bourgain's L^2 estimate of the maximal operator along analytic vector fields.

Acknowledgements

I would like to thank my supervisor Prof. Dr. Christoph Thiele for his enormous patience and help during my Ph.D. studies. As his first student in Bonn, I enjoyed many privileges. For example, in the first year of my Ph.D., I was Prof. Thiele's only student. He spent a considerably large amount of time and energy helping me out with my research, and that was the most wonderful and enjoyable time in the past three years. Every Monday afternoon I would knock on the door to his office, and carry out a discussion lasting a couple of hours. I would report to him on the progress I made and the difficulties I faced during the past week. Prof. Thiele would come up with a plethora of ideas in return. I would then implement his ideas one by one in order to greet the next meeting. Fortunately I managed to benefit from many great ideas from Prof. Thiele, and smoothly finished my first paper within one year. During my Ph.D. studies, there have been many other similar cases. If I were to list them all, I am afraid that the committee would not grant me any doctorate.

Moreover, I would also like to thank the other members from the Analysis and PDEs group. Particularly, I thank my Master's thesis adviser Prof. Dr. Herbert Koch for his strong recommendation to my Ph.D. application, thank Christian Zillinger for his patience in teaching a poor kid from a mountain village how to catch up with the pace of the modern high-tech times, thank Stefan Steinerberger and Angkana Rüland for teaching that kid how to work as a mathematician, thank Tobias Schottdorf, Diogo Oliveira e Silva, Mariusz Mirek, Gennady Uraltsev, Polona Durcik, Joris Roos and Lisa Onkes for making that kid's life in a foreign culture so gorgeous and colourful.

Last but not least, I would like to thank Yuexin Li and Diogo Oliveira e Silva for helping me to improve my English.

Contents

Chapter 1

Introduction and statement of the main results

The classical Lebesgue differentiation theorem states that for any $f \in L^1_{loc}(\mathbb{R}^n)$ we have

$$\lim_{\epsilon \to 0} \frac{1}{|B_\epsilon(x)|} \int_{B_\epsilon(x)} f(z)dz = f(x), \text{ a.e.} \tag{1.0.1}$$

with $B_\epsilon(x)$ denoting the ball of radius r centered at x. On the plane $\mathbb{R}^2$, instead of balls, taking averages over lower dimensional submanifolds like spheres or parabolas also appear naturally in harmonic analysis. A question I am interested in is the following: Suppose every point $x \in \mathbb{R}^2$ is assigned a unit vector $v(x)$, then given any $f \in L^2_{loc}(\mathbb{R}^2)$, does it hold true that

$$\lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} f(x + tv(x))dt = f(x), \text{ a.e. ?} \tag{1.0.2}$$

If the vector field is too "rough", the above pointwise convergence may fail almost everywhere. One such example can be constructed basing on the so-called Nikodym set. In 1927, Nikodym constructed a set $N \subset [0,1] \times [0,1]$ with the properties that $|N| = 1$, and that $\forall x \in N$, there exists a straight line passing through x which meets N only at x. The associated vector field v is defined as follows: at a point $x \in N$, let $v(x)$ be the unit vector parallel to the line that passes through x in the Nikodym set; at $x \notin N$, let $v(x)$ be any fixed unit vector, say $(1,0)$.

If we take the function f in (1.0.2) to be the characteristic function of the complement of the Nikodym set in the unit square, i.e.

$$f := \chi_{[0,1]\times[0,1]\setminus N}, \tag{1.0.3}$$

then we know that

$$f = 0, \text{ a.e. on } \mathbb{R}^2 \tag{1.0.4}$$

by the properties of the Nikodym set. However, for all $x \in N$, it is also easy to see that

$$\lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} f(x + tv(x))dt = 1. \tag{1.0.5}$$

Hence (1.0.2) fails almost everywhere on the unit square.

The above example tells us that in order for the pointwise convergence in (1.0.2) to be true, the vector field v has to have certain "regularity". Indeed it is a long standing conjecture in harmonic analysis (attributed to Zygmund in [21]) that (1.0.2) will hold true as long as we assume the vector field to be Lipschitz continuous. To state this conjecture quantitatively, it is convenient to introduce the following maximal function

$$M_{v,\epsilon_0} f(x) := \sup_{0<\epsilon\leq\epsilon_0} \frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} |f(x + tv(x))|dt, \tag{1.0.6}$$

where ϵ_0 is some positive constant depending on v. By a standard argument, (1.0.2) follows from the weak type estimate

$$\sup_{\lambda>0} \lambda^2 \left|\{x \in \mathbb{R}^2 : M_{v,\epsilon_0} f(x) > \lambda\}\right| \leq C\|f\|_2^2, \tag{1.0.7}$$

where $C > 0$ is some universal constant. Now we are ready to state

Conjecture 1 *([21]) There exists a universal constant $\kappa > 0$ such that for any unit Lipschitz vector field $v : \mathbb{R}^2 \to S^1$, the estimate* (1.0.7) *holds true with $\epsilon_0 := \kappa/\|v\|_{Lip}$.*

Instead of the maximal operator (1.0.6), it is also very natural to consider its singular integral variant

$$H_{v,\epsilon_0} f(x) := \int_{-\epsilon_0}^{\epsilon_0} f(x + tv(x))\frac{dt}{t} \tag{1.0.8}$$

as they share many features, in particular they have the same scaling and thus share the same potential L^p bounds. Indeed, the following conjecture was stated and studied by Lacey and Li in [21].

Conjecture 2 *([21]) There exists a universal constant $\kappa > 0$ such that for any unit Lipschitz vector field $v : \mathbb{R}^2 \to S^1$, the operator H_{v,ϵ_0} with $\epsilon_0 := \kappa/\|v\|_{Lip}$ satisfies the following weak type* $(2, 2)$ *estimate:*

$$\sup_{\lambda>0} \lambda^2 \left|\{x \in \mathbb{R}^2 : H_{v,\epsilon_0} f(x) > \lambda\}\right| \leq C\|f\|_2^2, \tag{1.0.9}$$

where C is a constant independent of f and v.

The constant ϵ_0 in the above Conjecture 1 and 2 is suggested by the isotropic scaling symmetry

$$x \to \lambda x. \tag{1.0.10}$$

Here we carry out the discussion by taking the example of the maximal operator. If we assume that there exists a constant $\kappa > 0$ such that for all v with $\|v\|_{Lip} \le 1$, the truncated maximal operator $M_{v,\kappa}$ satisfies the weak type estimate (1.0.7), then for an arbitrary Lipschitz vector field v, by choosing

$$\lambda = \frac{1}{\|v\|_{Lip}}, \tag{1.0.11}$$

and denoting $v_\lambda(\cdot) = v(\lambda \cdot)$, we obtain that

$$\|v_\lambda(\cdot)\|_{Lip} = 1. \tag{1.0.12}$$

Hence by our assumption, $M_{v_\lambda,\kappa}$ satisfies the weak type estimate (1.0.7). Next, if we denote $f_\lambda(\cdot) := f(\lambda \cdot)$, then by a simple change of variable, we observe that

$$\begin{aligned} M_{v_\lambda,\kappa} f_\lambda(x) &= \sup_{0<\epsilon\le\kappa} \frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon} |f(\lambda x + \lambda t v(\lambda x))| dt \\ &= \sup_{0<\epsilon\le\lambda\kappa} \frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon} |f(\lambda x + t v(\lambda x))| dt = M_{v,\lambda\kappa} f(\lambda x). \end{aligned} \tag{1.0.13}$$

Hence $M_{v,\lambda\kappa}$, which can also be written as $M_{v,\kappa/\|v\|_{Lip}}$, satisfies the same weak type estimate as (1.0.7).

So far we have seen that the truncation $\epsilon_0 = \kappa/\|v\|_{Lip}$ of the maximal operator (1.0.6) and the Hilbert transform (1.0.8) appears naturally for Lipschitz vector fields. Next we will state several motivations of conjecturing the Lipschitz regularity. Still, we will take the maximal operator as example.

1.1 Motivation for Lipschitz regularity

From change of variables. Let v be a unit vector field such that $\|v\|_{Lip} \le 1$. For a small enough κ and for any $0 < \epsilon \le \kappa$, we consider the average at the scale ϵ, which is defined by

$$A_{v,\epsilon} f(x) := \frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon} |f(x + t v(x))| dt. \tag{1.1.1}$$

Notice that

$$M_{v,\epsilon_0}f(x) = \sup_{0<\epsilon\leq\epsilon_0} A_{v,\epsilon}f(x). \tag{1.1.2}$$

For any $p \in (1,\infty)$, by Minkowski's inequality, we have

$$\left\| \frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon} |f(x+tv(x))|dt \right\|_p \leq \frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon} \|f(x+tv(x)\|_p\, dt. \tag{1.1.3}$$

For any fixed $t \in [-\epsilon,\epsilon]$, by the assumption that $\|v\|_{Lip} \leq 1$ and a simple change of variables, it is not difficult to see that

$$\|f(x+tv(x)\|_p \leq C\|f\|_p \tag{1.1.4}$$

for some constant $C > 0$. Hence we obtain that

$$\sup_{0<\epsilon\leq\epsilon_0} \|A_{v,\epsilon}f\|_p \leq C\|f\|_p. \tag{1.1.5}$$

Compared with (1.1.5), the estimate that Conjecture 1 aims at is

$$\left\| \sup_{0<\epsilon\leq\epsilon_0} A_{v,\epsilon}f \right\|_p \leq C\|f\|_p, \tag{1.1.6}$$

which is of a very similar form. Hence it is natural to conjecture that the above estimate (1.1.6) also holds true.

The Knapp example. For any given $\alpha \in (0,1)$, we will construct a counter example of C^α vector field v, such that the associated maximal operator M_{v,ϵ_0} is not bounded on L^p for any $p \leq 2$. Before constructing this example, we should make clear the dependence of the truncation ϵ_0 on the C^α norm of the vector fields. By a similar scaling argument to (1.0.10)-(1.0.13), ϵ_0 in the maximal operator M_{v,ϵ_0} should be determined by

$$\epsilon_0 := \kappa/\|v\|_{C^\alpha}^{1/\alpha}, \tag{1.1.7}$$

for some small positive constant κ. Here we use $\|\cdot\|_{C^\alpha}$ to denote the homogeneous Hölder norm, which is

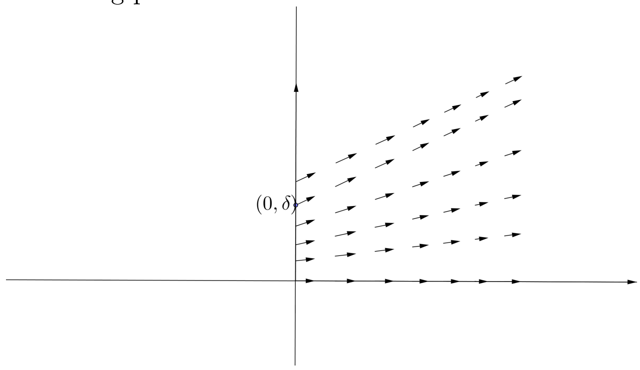
$$\|v\|_{C^\alpha} := \sup_{x\neq y} \frac{|v(x)-v(y)|}{|x-y|^\alpha}, \tag{1.1.8}$$

as we always consider unit vector fields.

Now we proceed with the detailed construction of the counter example. For some small $\delta > 0$, in the first quadrant of the plane, define the vector field v_δ by

$$v_\delta(x_1, x_2) := \begin{cases} (1, \delta) & \text{if } x_2 \geq \delta(x_1 + 1); \\ (1, \frac{x_2}{1+x_1}) & \text{else.} \end{cases} \tag{1.1.9}$$

The following picture illustrates how the above vector field looks like in the first quadrant:



When calculating the C^α norm of the vector field v_δ in the first quadrant, we observe that v_δ changes fastest along the Y-axis, hence we obtain that

$$\|v_\delta\|_{C^\alpha} = \sup_{0 \leq x_2 \leq \delta} \frac{x_2}{x_2^\alpha} = \delta^{1-\alpha}. \tag{1.1.10}$$

Next it is not difficult to extend the definition of v_δ to the whole plane satisfying the same bound as above. Hence the truncation ϵ_0 is given by

$$\epsilon_0 = \delta^{-\frac{1-\alpha}{\alpha}}. \tag{1.1.11}$$

Now we need to pick a "bad" function f_δ corresponding to the vector field v_δ. Denote

$$R_\delta := [-1-\delta, -1+\delta] \times [-\delta^2, \delta^2], \tag{1.1.12}$$

define

$$f_\delta := \chi_{R_\delta}. \tag{1.1.13}$$

The L^p norm of the function f_δ is

$$\|f_\delta\|_p = \delta^{\frac{3}{p}}. \tag{1.1.14}$$

Concerning the output of the operator M_{v_δ,ϵ_0}: at the point (x_1, x_2) with

$$0 \leq x_1 \leq \delta^{-\frac{1-\alpha}{\alpha}} \text{ and } 0 \leq x_2 \leq \delta(1 + x_1), \tag{1.1.15}$$

we have that

$$M_{v_\delta,\epsilon_0} f_\delta(x_1, x_2) \geq \frac{\delta}{1 + x_1}. \tag{1.1.16}$$

Hence

$$\begin{aligned}
\|M_{v_\delta,\epsilon_0} f_\delta\|_p^p &\geq \int_0^{\delta^{-\frac{1-\alpha}{\alpha}}} \int_0^{\delta(1+x_1)} \left(\frac{\delta}{1+x_1}\right)^p dx_2 dx_1 \\
&\geq \int_0^{\delta^{-\frac{1-\alpha}{\alpha}}} \delta^{p+1}(1 + x_1)^{-p+1} dx_1 \\
&\geq \begin{cases} C_{p,\alpha}\delta^{p+1}\delta^{-\frac{1-\alpha}{\alpha}\cdot(2-p)} & \text{for } p < 2. \\ C_{2,\alpha}\delta^3 \ln \frac{1}{\delta} & \text{for } p = 2. \end{cases}
\end{aligned} \tag{1.1.17}$$

Here for $p \leq 2$, $C_{p,\alpha} > 0$ is a constant depending on p and α. From the above estimate we conclude that for any $p \leq 2$, the maximal operator M_{v_δ,ϵ_0} is not bounded on L^p.

1.2 Partial progress towards Conjecture 1 and 2

In terms of regularity, the only known results concerning Conjecture 1 and 2 are for analytic vector fields. Bourgain [7] proved that for any analytic vector field v, there exists ϵ_0 depending on v such that M_{v,ϵ_0} is bounded on L^2. The L^p bounds for all $p > 1$ for both the maximal operator M_{v,ϵ_0} and the Hilbert transform H_{v,ϵ_0} were proved by Stein and Street [26]. Indeed, the results in [26] are much more general (including the multi-parameter case), but not in terms of regularity.

For smooth vector fields, Christ, Nagel, Stein and Wainger [9] proved, under some extra curvature conditions, that the associated maximal operator and singular integral operators are bounded on L^p for $p > 1$.

There is an interesting connection between the Hilbert transform along vector fields and Carleson's maximal operator, which was observed by Coifman and El Kohen, we review the discussion as presented in [5]. Consider the case of the one-variable vector fields, i.e. vector fields of the form $v(x_1, x_2) = (1, u(x_1))$ for some measurable function $u : \mathbb{R} \to \mathbb{R}$. Define the Hilbert transform along the vector field v without cut-off by

$$(H_v f)(x_1, x_2) := \int_{\mathbb{R}} f(x_1 - t, x_2 - u(x_1)t)\frac{dt}{t}. \tag{1.2.1}$$

Denoting by $\widehat{f}$ the partial Fourier transform in the second variable we obtain formally

$$\int f(x_1 - t, x_2 - u(x_1)t)\frac{dt}{t} \tag{1.2.2}$$

$$= \int e^{ix_2\xi_2} \int \widehat{f}(x_1 - t, \xi_2)e^{-iu(x_1)t\xi_2}\frac{dt}{t}d\xi_2.$$

By the Plancherel theorem, the L^2 norm of this expression in the x_1 and x_2 variables is the same as the L^2 norm in the variables x_1 and ξ_2 of the expression

$$\int \widehat{f}(x_1 - t, \xi_2)e^{iu(x_1)t\xi_2}\frac{dt}{t}.$$

For each fixed ξ_2, we recognise this to essentially be the linearisation of Carleson's maximal operator

$$(Cf)(x) := \sup_{N\in\mathbb{R}} \left| \int_{\mathbb{R}} f(x - t)e^{iNt}\frac{dt}{t} \right|. \tag{1.2.3}$$

The use of Plancherel's theorem makes this simple argument work only in L^2. To go beyond L^2, we need to replace the Fourier transform by a Littlewood-Paley decomposition. Lacey and Li [20], exploiting the connection between the Hilbert transform along vector fields and Lacey and Thiele's proof for the boundedness of the bilinear Hilbert transform [23] [24] and Carleson's maximal operator [25], proved that for any measurable vector field v, the operator $H_v P_k$, which is the composition of the Hilbert transform along v with a Littlewood-Paley projection operator P_k for some fixed k, maps L^2 to weak L^2, and L^p to L^p for $p > 2$, uniformly in k. Moreover, conditioning on the boundedness of what they called the Lipschitz-Kakeya maximal operator, Lacey and Li [20] also proved that for any $C^{1+\alpha}$ vector field v with $\alpha > 0$, the operator H_{v,ϵ_0} is bounded on L^2 for some properly chosen ϵ_0.

Afterwards, Bateman verified the boundedness of the Lipschitz-Kakeya maximal operator for the one-variable vector fields in [3]. On that basis Bateman [4], Bateman and Thiele [5] proved the following

Theorem 1.1 *([4], [5]) For a one variable vector field v of the form $v(x_1, x_2) = (1, u(x_1))$ for some measurable function u, the associated Hilbert transform H_v defined as in* (1.2.1) *is bounded on L^p*

for $p \in (3/2, \infty)$.

Remark 1.2 *Here we make an observation that Bateman and Thiele's result in [5] holds true for all Hörmander-Mihlin kernels. Let* K *be a Calderon-Zygmund kernel with* $m := \check{K}$ *satisfying the Hörmander-Mihlin condition*

$$|\partial^\beta m(\xi)| \leq \frac{C_\beta}{|\xi|^\beta}, \forall \xi \in \mathbb{R} \setminus \{0\} \tag{1.2.4}$$

for sufficiently many $\beta \in \mathbb{N}$, *where* $C_\beta > 0$ *is a constant depending only on* β. *For a one-variable vector field* $v(x_1, x_2) = (1, u(x_1))$ *for some measurable function* u, *similar to* (1.2.1), *define the associated singular integral with kernel* K *along* v *by*

$$H_v^K f(x_1, x_2) := \int_{\mathbb{R}} f(x_1 - t, x_2 - u(x_1)t) K(t) dt. \tag{1.2.5}$$

Then for all $p > 3/2$, *we claim that*

$$\|H_v^K f\|_p \leq C_{p,K} \|f\|_p, \tag{1.2.6}$$

with a constant $C_{p,K} > 0$ *depending only on* p *and the kernel* K. *The estimate* (1.2.6) *follows from Bateman and Thiele's proof of Theorem 1.1 and the anisotropic scaling*

$$x_1 \to x_1, x_2 \to \lambda x_2. \tag{1.2.7}$$

If we denote by Γ *the cone* $\{(x_1, x_2) : |x_1| \leq |x_2|\}$, $\mathbb{1}_\Gamma$ *the indicator function of the cone* Γ, *and* Π_Γ *the frequency projection operator on the cone* Γ, *which is*

$$\Pi_\Gamma f := \mathcal{F}^{-1}(\mathbb{1}_\Gamma \cdot \mathcal{F} f), \tag{1.2.8}$$

where $\mathcal{F}$ *denotes the Fourier transform and* $\mathcal{F}^{-1}$ *its inverse, then what Bateman and Thiele have proven in [5] is*

$$\|H_v^K \Pi_\Gamma f\|_p \leq C_{p,K} \|\Pi_\Gamma f\|_p, \tag{1.2.9}$$

under the assumption that $\|u\|_\infty \leq 1$. *Notice that* (1.2.7) *and* (1.2.9) *imply that*

$$\|H_v^K \Pi_{\Gamma_\lambda} f\|_p \leq C_{p,K} \|\Pi_{\Gamma_\lambda} f\|_p, \tag{1.2.10}$$

for all $\lambda \geq 1$, *where* Γ_λ *denotes the cone* $\{(x_1, x_2) : |x_1| \leq \lambda |x_2|\}$. *Hence by a limiting argument, we* *obtain*

$$\|H_v^K f\|_p \leq C_{p,K}\|f\|_p, \tag{1.2.11}$$

under the assumption that $\|u\|_\infty \leq 1$. *Again by another limiting argument, we can get rid of the restriction that* $\|u\|_\infty \leq 1$ *and obtain the estimate* (1.2.6).

Out of the same reasoning as in (1.2.1)-(1.2.3), *the* L^2 *boundedness in* (1.2.6) *is equivalent to the* L^2 *boundedness of Li and Muscalu's generalised Carleson's maximal operator in [22]:*

$$\sup_{N\in\mathbb{R}}\left|\int_{\mathbb{R}} f(x-t)e^{iNt}K(t)dt\right|, \tag{1.2.12}$$

where K *is a kernel satisfying the Hörmander-Mihlin condition* (1.2.4).

1.3 Statement of the main results

The first result I have obtained concerns a perturbation of Bateman and Thiele's result [5] in the critical Lipschitz regularity. This result is contained in my preprints [14] and [15].

Theorem 1.3 ([14], [15]) *For any measurable vector field* v_0 *satisfying the following two conditions*

i) there exists a bi-Lipschitz map $g_0 : \mathbb{R}^2 \to \mathbb{R}^2$ *s.t. for all* $x_1 \in \mathbb{R}$

$$v_0(g_0(x_1, x_2)) \text{ is constant in } x_2, \tag{1.3.1}$$

ii) there exists $d_0 > 0$ *s.t. for all* $x_1 \in \mathbb{R}$

$$\angle(\partial_2 g_0(x_1,\cdot), \pm v_0(g_0(x_1,\cdot))) \geq d_0 \text{ a.e. on } \mathbb{R}, \tag{1.3.2}$$

the associated Hilbert transform, which is defined by

$$H_{v_0}f(x) := \int_{\mathbb{R}} f(x - tv_0(x))\frac{dt}{t}, \tag{1.3.3}$$

is bounded on L^p *for all* $p > 3/2$. *Moreover, the operator norm blows up when* $d_0 \to 0$.

Remark 1.4 *By taking* g_0 *to be the identity map and applying the anisotropic scaling* (1.2.7) *correspondingly, we recover the result of Bateman and Thiele [5].*

Remark 1.5 *To our knowledge, this is the first result in the context of the Hilbert transform along vector fields with a Lipschitz regularity in its hypothesis. Indeed, a structure theorem for Lipschitz functions by Azzam and Schul [2] states that any Lipschitz function* $u_0 : \mathbb{R}^2 \to \mathbb{R}$ *(any Lipschitz unit vector field* v_0 *in our case) can be precomposed with a bi-Lipschitz function* $g_0 : \mathbb{R}^2 \to \mathbb{R}^2$ *such that* $u_0 \circ g_0$*, when restricted to a "large" portion of its domain, is Lipschitz in the first coordinate and constant in the second coordinate.*

The proof of Theorem 1.3 does not work for the maximal operator (1.0.6). Hence as part of the enterprise to understand the difference between the singular integral operator (1.0.8) and the maximal operator (1.0.6), we started to look at Bourgain's result [7] on the L^2 bounds of the maximal operator along analytic vector fields. In the meantime, we observed that some of the techniques used to prove Theorem 1.3 could be applied to provide a geometric proof of Bourgain's result mentioned above, which is

Theorem 1.6 ([7]) *Let* v *be an analytic vector field on a bounded set* Ω*. Then for* ϵ_0 *chosen small enough, the associated maximal operator* M_{v,ϵ_0} *defined in* (1.0.6) *is bounded on* L^2*.*

The main tool that is used in the proof of Theorem 1.6 is the time-frequency decomposition initiated by Lacey and Li in the setting of the Hilbert transform along vector fields in [20] and [21]. However, the proof is free of the time-frequency analysis techniques.

Notations: Throughout this thesis, we will write $x \ll y$ to mean that $x \leq y/10$, $x \lesssim y$ to mean that there exists a constant C s.t. $x \leq Cy$, and $x \sim y$ to mean that $x \lesssim y$ and $y \lesssim x$. $\mathbb{1}_E$ will always denote the characteristic function of the set E.

Chapter 2

Hilbert transform along measurable vector fields constant on Lipschitz curves: L^2 boundedness

In this chapter, we will present the proof of Theorem 1.3 for the case $p = 2$. The content of this chapter is essentially contained in my preprint [14]. Before starting the proof, we would like to look at Bateman and Thiele's result in [5] (Theorem 1.1) from another point of view.

On $\mathbb{R}^2$, a direction is given by vector $v = (1, u)$, where $u \in \mathbb{R}$. Consider the directional Hilbert transform in the plane defined for a fixed direction $v = (1, u)$ as

$$H_v f(x, y) := \int_{\mathbb{R}} f(x - t, y - ut) \frac{dt}{t} \tag{2.0.1}$$

for any test function f. Here we use (x, y) instead of (x_1, x_2) to denote one point in $\mathbb{R}^2$ as we want to emphasis the bi-parameter structure of this operator. The bi-parameter structure arises since the kernel is a tensor product between a Hilbert kernel in direction v and a Dirac delta distribution in the perpendicular direction. By the dilation symmetry, the length of the vector v is irrelevant for the value of H_v, which explains our normalisation of the first component.

By an application of Fubini's theorem and the L^p bounds for the classical Hilbert transform one obtains a priori L^p bounds for H_v. On the other hand, the corresponding maximal operator $\sup_u |H_v f(x, y)|$ for varying directions is well known to not satisfy any a priori L^p bounds, see the work of Karagulyan [17].

What Bateman and Thiele proved in [5] is

$$\left\| \int_{\mathbb{R}} f(x - t, y - u(x)t) \frac{dt}{t} \right\|_p \lesssim \|f\|_p, \tag{2.0.2}$$

where $u : \mathbb{R} \to \mathbb{R}$ is any measurable function and the constant depends only on p for $p > 3/2$. The estimate (2.0.2) can also be interpreted as

$$\| \sup_{u \in \mathbb{R}} \|H_v f(x,y)\|_{L^p(y)} \|_{L^p(x)} \lesssim \|f\|_p, \tag{2.0.3}$$

as it is just a linearisation of the maximal operator in (2.0.3). The function $u : \mathbb{R} \to \mathbb{R}$ in (2.0.2) will be called the linearising function. Note that the supremum falls between the computation of the norm in y and in x, compared to being completely inside or outside as in our previous remarks. The estimate (2.0.3) also highlights the bi-parameter structure of the directional Hilbert transform.

In the following, we will first formulate a generalisation of Bateman and Thiele's result in the form of (2.0.3), and then derive Theorem 1.3 as a corollary. To formulate such a result, we perturb (2.0.3) by a bi-Lipschitz horizontal distortion, that is

$$(x, y) \to (g(x, y), y) \tag{2.0.4}$$

with

$$(x' - x)/a_0 \leq g(x', y) - g(x, y)) \leq a_0(x' - x) \tag{2.0.5}$$

for every $x < x'$ and every y, such that the transformation (2.0.4) maps vertical lines to near vertical Lipschitz curves:

$$|g(x, y) - g(x, y')| \leq b_0|y' - y| \tag{2.0.6}$$

for all x, y, y'. These two conditions can be rephrased as

$$1/a_0 \leq \partial_1 g \leq a_0 \text{ and } |\partial_2 g| \leq b_0 \text{ a.e.} \tag{2.0.7}$$

Under these assumptions, L^p norms are distorted boundedly under the transformation (2.0.4). Namely, (2.0.5) implies for every y that

$$a_0^{-1}\|f(x, y)\|^p_{L^p(x)} \leq \|f(g(x, y), y)\|^p_{L^p(x)} \leq a_0\|f(x, y)\|^p_{L^p(x)} \tag{2.0.8}$$

and we may integrate this in y direction to obtain equivalence of L^p norms in the plane. Hence the change of measure is not the main point of the following theorem, but rather the effect of the transformation on the linearising function u, which is now constant along the family of Lipschitz curves which are the images of the lines $x = x_0$ under the map (2.0.4).

Theorem 2.1 *Let $g : \mathbb{R}^2 \to \mathbb{R}$ satisfy assumption (2.0.5) for some a_0 and assumption (2.0.6) for some b_0. Then for any $c_0 \in (0,1)$, we have*

$$\| \sup_{|u| \le c_0/b_0} \|H_v f(g(x,y),y)\|_{L^2(y)}\|_{L^2(x)} \le C\|f\|_2. \tag{2.0.9}$$

Here C is a constant depending only on a_0 and c_0.

Remark 2.2 *The constant C is independent of b_0 due to the anisotropic scaling symmetry $x \to x, y \to \lambda y$.*

The case $p = 2$ in Theorem 1.3 will be derived as a corollary of the above Theorem 2.1. We will present the reduction in the following Section 2.1. Before doing that, we first comment on the proof of Theorem 2.1.

To use the assumption that the linearising function v of (2.0.9) is constant along Lipschitz curves, we apply an adapted Littlewood-Paley theory along the level lines of v. This is a refinement of the analysis of Coifman and El Kohen who use a Fourier transform in the y variable and the analysis of Bateman and Thiele who use a classical Littlewood-Paley theory in the y variable. This adapted Littlewood-Paley theory is the main novelty of the proof. It is in the spirit of prior work on the Cauchy integral on Lipschitz curves, for example [10] , but it differs from this classical theme in that it is more of bi-parameter type as it is governed by a whole fibration into Lipschitz curves. We crucially use Jones' beta numbers as a tool to control the adapted Littlewood-Paley theory. To our knowledge this is also the first use of Jones' beta numbers in the context of the directional Hilbert transform.

In this chapter we focus on the case L^2, since our goal here is to highlight the use of the adapted Littlewood-Paley theory and Jones' beta numbers in the technically most simple case.

While Coifman and El Kohen use the difficult bounds on Carleson's operator as a black box, Bateman and Thiele have to unravel this black box following the work of Lacey and Li [20]-[21] and use time-frequency analysis to prove bounds for a suitable generalisation of Carleson's operator. Luckily, in the proof of Theorem 1.3 we do not have to delve into time-frequency analysis as we can largely recycle the work of Bateman and Thiele for this aspect of the argument.

An upper bound such as $|u| \le c_0/b_0$ is necessary in our theorem. By a limiting argument we may recover the theorem of Bateman and Thiele, using the scaling to tighten the Lipschitz constant b_0 at the same time as relaxing the condition $|u| \le c_0/b_0$.

An interesting open question remains whether the same holds true for $c_0 = 1$. We do not know of a soft argument to achieve this relaxation. Our estimate of the norms become unbounded as c_0 approaches 1. This question suggests itself for further study.

Outline of chapter: in Section 2.1 we will prove the L^2 bounds in Theorem 1.3 by reducing it to Theorem 2.1. The reduction will also be used later in the proof of Theorem 2.1.

In Section 2.2 we will state the strategy of the proof for Theorem 2.1. As it appears that our result is a Lipschitz perturbation of the one by Bateman and Thiele, this turns out also to be the case for the proof: if we denote by P_k a Littlewood-Paley operator in the y-variable, the main observation in Bateman and Thiele's proof is that H_v (given by (1.2.1)) commutes with P_k. In our case, this is no longer true. However, we can make use of an adapted version of the Littlewood-Paley projection operator $\tilde{P}_k$ (see Definition 2.5) to partially recover the orthogonality. We split the linearised operator H_v of the left hand side of (2.0.9) into a main term and a commutator term

$$\sum_{k\in\mathbb{Z}} H_v P_k(f) = \sum_{k\in\mathbb{Z}} (H_v P_k(f) - \tilde{P}_k H_v P_k(f) + \tilde{P}_k H_v P_k(f)). \tag{2.0.10}$$

The boundedness of the main term $\sum_{k\in\mathbb{Z}} \tilde{P}_k H_v P_k(f)$ is essentially due to Lacey and Li [20], with conditionality on certain maximal operator estimate. In Section 2.3 we modify Bateman's argument in [3] and [4] to the case of vector fields constant on Lipschitz curves and remove the conditionality on that maximal operator.

The main novelty is the boundedness of the commutator term

$$\sum_{k\in\mathbb{Z}} (H_v P_k(f) - \tilde{P}_k H_v P_k(f)), \tag{2.0.11}$$

which will be presented in Section 2.4. To achieve this, we will view Lipschitz curves as perturbations of straight lines and use Jones' beta number condition for Lipschitz curves and the Carleson embedding theorem to control the commutator. Here we shall emphasis again that the commutator estimate is free of time-frequency analysis.

2.1 Proof of the L^2 bounds in Theorem 1.3 by reducing to Theorem 2.1

In this section we prove the L^2 bounds in Theorem 1.3, by reducing it to Theorem 2.1. The reduction is based on a cutting and pasting argument. Some parts of the reduction will also be used in the proof of Theorem 2.1 in the rest of the chapter.

We first divide the unit circle S^1 into N arcs of equal length, with the angle of each arc being $2\pi/N$. Choose

$$N > 6\pi/d_0 \tag{2.1.1}$$

s.t. $2\pi/N < d_0/3$. Denote these arcs as $\Omega_1, \Omega_2, ..., \Omega_N$. For each Ω_i, define

$$H_{v_0,\Omega_i} f(x, y) := \begin{cases} H_{v_0} f(x, y) & \text{if } v_0(x, y) \in \Omega_i \\ 0 & \text{else} \end{cases}$$

where $H_{v_0} f$ is given by (1.3.3). If we were able to prove that $\|H_{v_0,\Omega_i}\|_{2\to 2}$ is bounded by a constant C which is independent of $i \in \{1, 2, ..., N\}$, then we conclude that

$$\|H_{v_0}\|_{2\to 2} \leq CN(d_0). \tag{2.1.2}$$

Now fix one Ω_i, we want to show the boundedness of H_{v_0,Ω_i}. Choose a new coordinate such that the x-axis passes through Ω_i and bisects it. Then all the vectors in Ω_i form an angle less than $d_0/6$ with the x-axis. As we assume that

$$\angle(\partial_2 g_0, \pm v_0(g_0)) \geq d_0 > 0, \tag{2.1.3}$$

we see that the vector $\partial_2 g_0$ forms an angle less than $\frac{\pi - d_0}{2}$ with the y-axis.

Renormalise the unit vector v_0 such that the first component is 1, i.e. write $v_0 = (1, u_0)$, then by the fact that v_0 forms an angle less than $d_0/6$ with the x-axis, we obtain

$$|u_0| \leq \tan(d_0/6). \tag{2.1.4}$$

Next we construct the Lipschitz function g in Theorem 2.1 from the bi-Lipschitz map g_0, and the coordinate we will use here is still the one associated to Ω_i as above. Under this linear change of variables, we know that g_0 is still bi-Lipschitz. We renormalise the bi-Lipschitz map in such a way that

$$g_0(x,0) = (x,0), \forall x \in \mathbb{R}. \tag{2.1.5}$$

Fix $x \in \mathbb{R}$, the map g_0, when restricted on the vertical line $\{(x,y) : y \in \mathbb{R}\}$, is still bi-Lipschitz. We denote by Γ_x the image of this bi-Lipschitz map, i.e.

$$\Gamma_x := \{g_0(x,y) : y \in \mathbb{R}\}. \tag{2.1.6}$$

Define the function g by the following relation

$$(g(x,y), y) = g_0(x, y\prime), \tag{2.1.7}$$

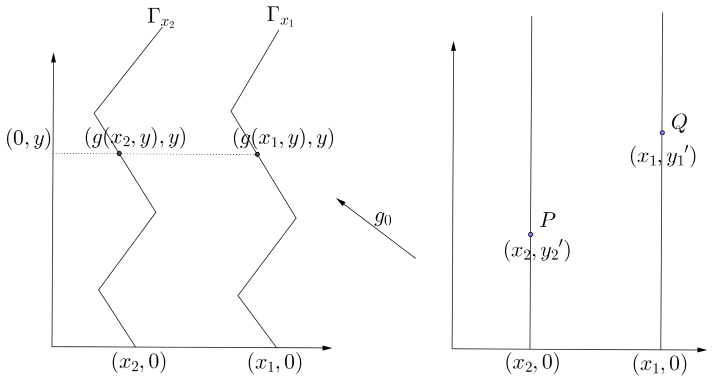
for some $y\prime$. By the fact that g_0 is bi-Lipschitz, we know that such $y\prime$ exists and is unique.

From the above construction and the fact that $\partial_2 g_0$ forms an angle less than $\frac{\pi - d_0}{2}$ with the y-axis, we see easily that

$$|g(x,y_1) - g(x,y_2)| \leq \cot(d_0/2)|y_1 - y_2|, \forall x, y_1, y_2 \in \mathbb{R}. \tag{2.1.8}$$

Hence what is left is to show that condition (2.0.5) is also satisfied with a constant a_0 depending only on d_0 and the bi-Lipschitz constant of g_0. One side of the equivalence $(x_1 - x_2)/a_0 \leq g(x_1,y) - g(x_2,y)$ is quite clear from the picture below: the bi-Lipschitz map g_0 sends the points P, Q to $(g(x_1,y),y), (g(x_2,y),y)$ separately, then by definition of bi-Lipschitz map, there exists constant a_0 s.t.

$$g(x_1,y) - g(x_2,y) \geq \frac{1}{a_0}|P - Q| \geq \frac{1}{a_0}(x_1 - x_2). \tag{2.1.9}$$



For the other side, we argue by contradiction. If for any $M \in \mathbb{N}$ large, there exists $x_1, x_2, y \in \mathbb{R}$ s.t.

$$g(x_1, y) - g(x_2, y) \geq M(x_1 - x_2), \tag{2.1.10}$$

then together with (2.1.8), this implies that

$$dist(K, \Gamma_{x_1}) \geq M \sin(d_0/2)(x_1 - x_2). \tag{2.1.11}$$

But this is not allowed as by the definition of the bi-Lipschitz map g_0 and the Lipschitz function g, $dist(K, \Gamma_{x_1})$ must be comparable to $|x_1 - x_2|$.

So far, we have verified all the conditions in Theorem 2.1 with

$$b_0 = \cot(d_0/2) \text{ and } c_0 = \tan(d_0/6)/\cot(d_0/2) < 1. \tag{2.1.12}$$

Hence we can apply Theorem 2.1 to obtain the boundedness of H_{v_0,Ω_i}.

In the end, as claimed in Theorem 1.3, we still need to show that the operator norm in L^p ($\forall p > 1$) blows up without the assumption that $d_0 > 0$. For the range $p \leq 2$, the counter example is simply a Knapp example: let $B_1(0)$ denote the ball of radius one centred at origin, take the function $f(x) = \mathbb{1}_{B_1(0)}(x)$, let Γ be the upper cone which forms an angle less than $\frac{\pi}{4}$ with the vertical axis. First define the vector field $v_0(x) = \frac{x}{|x|}$ for $x \in \Gamma \setminus B_1(0)$, then extend the definition to the whole plane properly such that v_0 satisfies the condition (1.3.1). It is then easy to see that

$$|H_{v_0} f(x)| \sim \frac{1}{|x|}, \forall x \in \Gamma \setminus B_1(0), \tag{2.1.13}$$

which does not belong to $L^p(\mathbb{R}^2)$ for $p \leq 2$. For the range $p > 2$, the counter example is given by the standard Besicovitch-Kakeya set construction, which can be found in Page 1022 [4] and Page 7 [20].

2.2 Strategy of the proof of Theorem 2.1

If we linearise the maximal operator in Theorem 2.1, what we need to prove turns to be the following

$$\| \int_{\mathbb{R}} f(g(x,y) - t, y - tu(x)) dt/t \|_2 \lesssim \|f\|_2, \tag{2.2.1}$$

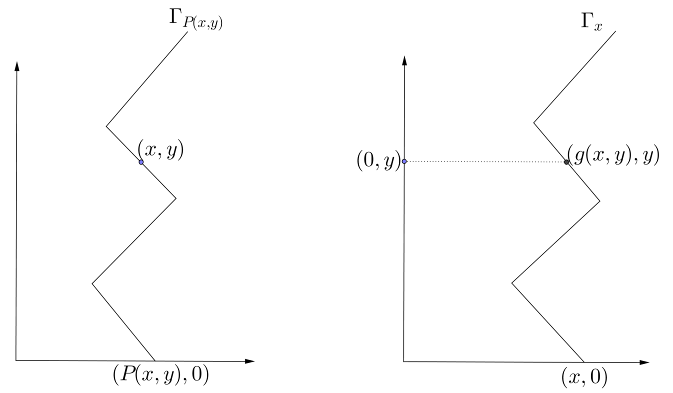
where $u : \mathbb{R} \to \mathbb{R}$ is a measurable function such that $\|u\|_\infty \leq c_0/b_0$. The change of coordinates

$$(x, y) \to (g(x,y), y) \tag{2.2.2}$$

in (2.0.4) also changes the measure on the plane. However, we still want to use the original Lebesgue measure for the Littlewood-Paley decomposition. Hence we invert (2.0.4) and denote the inversion by

$$(x, y) \to (P(x,y), y), \tag{2.2.3}$$

where "P" stands for "projection". The following picture illustrates why we call the map (2.2.3) a projection:



The change of coordinates in (2.2.3) turns the estimate (2.2.1) into the following equivalent form

$$\| \int_{\mathbb{R}} f(x - t, y - tu(P(x,y))) dt/t \|_2 \lesssim \|f\|_2. \tag{2.2.4}$$

Moreover, we will denote

$$H_v f(x,y) := \int_{\mathbb{R}} f(x-t, y-tu(P(x,y)))dt/t. \tag{2.2.5}$$

In the rest of the paper, we want to make the convention that whenever H_v appears, it denotes the Hilbert transform along the vector field $v(x,y) = (1, u(P(x,y)))$, which is the above (2.2.5), to distinguish it from the various H_v that have appeared previously in this chapter.

To prove the above estimate, we first make several reductions: by the anisotropic scaling

$$x \to x, y \to \lambda y, \tag{2.2.6}$$

we can w.l.o.g. assume that $b_0 = 10^{-2}$. By a similar cutting and pasting argument to that in the proof of Theorem 1.3, we can assume that $c_0 \ll 10^{-2}$, i.e. the vector field v is of the form $(1, u)$ with $|u| \ll 1$.

Now we start the proof. It was already observed in Bateman [4] (Page 1024) that under the assumption $|u| \ll 1$, we can w.l.o.g. assume that supp $\hat{f}$ lies in a two-ended cone which forms an angle less than $\pi/4$ with the vertical axis, as for functions f with frequency supported on the cone near the horizontal axis, we have that

$$H_v f(x,y) = H_{(1,0)} f(x,y), \tag{2.2.7}$$

which is the Hilbert transform along the constant vector field $(1,0)$. But $H_{(1,0)}$ is bounded by Fubini's theorem and the L^2 boundedness of the Hilbert transform.

For the frequencies outside the cone near the horizontal axis, the proof consists of two steps. In the first step we will prove the boundedness of H_v when acting on functions with frequency supported in one single annulus. To be precise, let Γ be the cone which forms an angle less than $\pi/4$ with the vertical axis, Π_Γ be the projection operator on Γ, i.e.

$$\Pi_\Gamma f := \mathcal{F}^{-1}(\mathbb{1}_\Gamma \cdot \mathcal{F} f), \tag{2.2.8}$$

where $\mathcal{F}$ stands for the Fourier transform and $\mathcal{F}^{-1}$ the inverse transform. Let P_k be the k-th Littlewood-Paley projection operator in the vertical direction, namely if we denote by ψ_0 is a smooth function with support on $[-5/2, -1/2] \cup [1/2, 5/2]$ such that

$$\sum_{k \in \mathbb{Z}} \psi_k(t) = 1, \forall t \neq 0, \tag{2.2.9}$$

with

$$\psi_k(t) := \psi_0(2^{-k} t), \tag{2.2.10}$$

then

$$P_k f(x, y) := \int_{\mathbb{R}} f(x, y - y') \check{\psi}_k(y') dy'. \tag{2.2.11}$$

As we are always concerned with the frequency in Γ, later for simplicity we will just write P_k instead of $P_k \circ \Pi_\Gamma$ for short. Then what we will prove first is

Proposition 2.3 *Under the same assumptions as in Theorem 2.1, we have for $p \in (1, \infty)$ that*

$$\|H_v P_k(f)\|_p \lesssim \|P_k(f)\|_p, \tag{2.2.12}$$

with the constant being independent of $k \in \mathbb{Z}$.

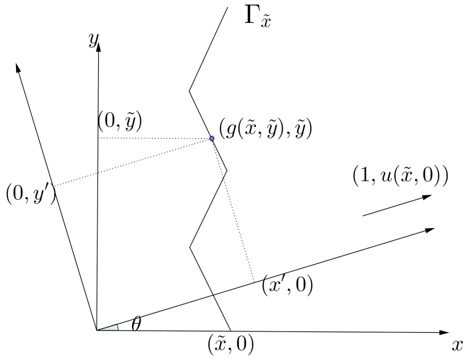
In order to prove the boundedness of H_v, we need to put all the frequency pieces together. In the case of $C^{1+\alpha}$ vector fields for any $\alpha > 0$, Lacey and Li's idea in [21] is to prove the almost orthogonality between different frequency annuli. In the case where the vector field is constant along vertical lines, an important observation in the paper of Bateman and Thiele is that H_v and P_k commute, which then makes it possible to apply a Littlewood-Paley square function estimate.

In our case Bateman and Thiele's observation is no longer true. We need to take into account that the vector field is constant along Lipschitz curves, which gives rise to an adapted Littlewood-Paley projection operator (the following Definition 2.5).

Before defining this operator, we first need to make some preparation. Fix one $\tilde{x} \in \mathbb{R}$, take the curve $\Gamma_{\tilde{x}}$ which passes through $(\tilde{x}, 0)$, recall that $\Gamma_{\tilde{x}}$ is given by the set $\{(g(\tilde{x}, \tilde{y}), \tilde{y}) : \tilde{y} \in \mathbb{R}\}$, where g is the Lipschitz function in Theorem 2.1. By the definition of the operator H_v we know that the vector field v is equal to the constant vector $(1, u(\tilde{x}))$ along $\Gamma_{\tilde{x}}$. Change the coordinate s.t. the horizontal x'-axis is parallel to $(1, u(\tilde{x}))$. The following lemma says that in the new coordinate, the curve $\Gamma_{\tilde{x}}$ can still be realised as the graph of a Lipschitz function.

Lemma 2.4 *For any fixed $\tilde{x} \in \mathbb{R}$, there exists a Lipschitz function $x' = g_{\tilde{x}}(y')$ s.t. $\Gamma_{\tilde{x}}$ can be re-parametrised as $\{(g_{\tilde{x}}(y'), y') : y' \in \mathbb{R}\}$. Moreover, we have that $\|g_{\tilde{x}}\|_{Lip} \leq \frac{1+b_0}{1-b_0}$, where b_0 is the* *constant in Theorem 2.1.*

Proof of Lemma 2.4: denote by θ the angle between the vector $(1, u(\tilde{x}))$ and the x-axis as in the picture below.



The new coordinate of the point $(g(\tilde{x}, \tilde{y}), \tilde{y})$ will be given by

$$(x', y') = (\tilde{y} \sin\theta + g(\tilde{x}, \tilde{y}) \frac{1 + \sin^2\theta}{\cos\theta}, \tilde{y}\cos\theta - g(\tilde{x}, \tilde{y})\sin\theta). \tag{2.2.13}$$

Look at the identity for the second component

$$y' = \tilde{y}\cos\theta - g(\tilde{x}, \tilde{y})\sin\theta, \tag{2.2.14}$$

we want to solve $\tilde{y}$ by y' by using the implicit function theorem. As

$$\frac{dy'}{d\tilde{y}} = \cos\theta - \frac{\partial g}{\partial \tilde{y}} \sin\theta, \tag{2.2.15}$$

by the fact that $|u| \ll 1$ and $|\frac{\partial g}{\partial \tilde{y}}| \leq b_0 \leq 10^{-2}$, we obtain that

$$\frac{1 - b_0}{\sqrt{2}} \leq \frac{dy'}{d\tilde{y}} \leq \frac{1 + b_0}{\sqrt{2}}, \tag{2.2.16}$$

from which it is clear that the implicit function theorem is applicable.

After solving $\tilde{y}$ by y', we just need to substitute $\tilde{y}$ into the identity for the first component in (2.2.13), which is

$$x' = \tilde{y}\sin\theta + g(\tilde{x}, \tilde{y}) \frac{1 + \sin^2\theta}{\cos\theta}, \tag{2.2.17}$$

to get an implicit expression of x' in terms of y', which we will denote as $x' = g_{\tilde{x}}(y')$.

To estimate the Lipschitz norm of the function $g_{\tilde{x}}$, we just need to observe that when doing the above change of variables, we have rotated the axis by an angle θ which satisfies $|\theta| \leq \pi/4$. Together with the fact that $|\frac{\partial g}{\partial \tilde{y}}| \leq b_0$, we can then derive that

$$|\frac{\partial g_{\tilde{x}}}{\partial y'}| \leq \frac{1+b_0}{1-b_0}, \tag{2.2.18}$$

which finishes the proof of Lemma 2.4. □

Definition 2.5 *(adapted Littlewood-Paley projection) Select a Schwartz function ψ_0 with support on $[\frac{1}{2}, \frac{5}{2}] \cup [-\frac{5}{2}, -\frac{1}{2}]$ such that*

$$\sum_{k\in\mathbb{Z}} \psi_0(2^{-k}t) = 1, \forall t \neq 0. \tag{2.2.19}$$

For $f : \mathbb{R}^2 \to \mathbb{R}$, for every fixed $\tilde{x} \in \mathbb{R}$, define the adapted (one dimensional) Littlewood-Paley projection on $\Gamma_{\tilde{x}}$ by

$$\tilde{P}_k(f)(x', y') := \int_{\mathbb{R}} f(g_{\tilde{x}}(z), z)\check{\psi}_k(y'-z)dz = P_k(\tilde{f})(y'), \tag{2.2.20}$$

where $(x', y') = (g_{\tilde{x}}(y'), y')$ denotes one point in $\Gamma_{\tilde{x}}$, $\psi_k(\cdot) := \psi_0(2^{-k}\cdot)$ and we use $\tilde{f}(\cdot)$ to denote the function $f(g_{\tilde{x}}(\cdot), \cdot)$, and P_k the one dimensional Littlewood-Paley projection operator.

Now it is instructive to regard the Lipschitz curves as perturbation of the straight lines, or equivalently, to think that H_vP_kf still has frequency supported near the k-th frequency band, which has already been used by Lacey and Li in their almost orthogonality estimate for $C^{1+\alpha}$ vector fields in [21]. We then subtract the term $\tilde{P}_kH_vP_k(f)$ from $H_vP_k(f)$, and estimate the commutator.

To be precise, we first write

$$\sum_k H_vP_k(f) = \sum_k (H_vP_k(f) - \tilde{P}_kH_vP_k(f) + \tilde{P}_kH_vP_k(f)), \tag{2.2.21}$$

then by the triangle inequality, we have

$$\|\sum_k H_vP_k(f)\|_2 \lesssim \|\sum_k (H_vP_k(f) - \tilde{P}_kH_vP_k(f))\|_2 + \|\sum_k \tilde{P}_kH_vP_k(f)\|_2. \tag{2.2.22}$$

We call the second term the main term, and the first term the commutator term. The L^2 boundedness of the main term will follow from orthogonality argument, which is the following adapted Littlewood-Paley theorem.

Lemma 2.6 *For $p \in (1, +\infty)$, we have the following variants of the Littlewood-Paley estimates:*

$$\|(\sum_{k\in\mathbb{Z}} |\tilde{P}_k(f)|^2)^{1/2}\|_p \sim \|f\|_p, \tag{2.2.23}$$

$$\|(\sum_{k\in\mathbb{Z}} |\tilde{P}_k^*(f)|^2)^{1/2}\|_p \sim \|f\|_p, \tag{2.2.24}$$

with constants depending only on a_0.

Proof of Lemma 2.6: In the above equation (2.0.8), we have already explained the following co-area formula:

$$\int_{\mathbb{R}^2} |f(x,y)|dxdy \sim \int_{\mathbb{R}} [\int_{\Gamma_{\tilde{x}}} |f| ds_{\tilde{x}}] d\tilde{x}. \tag{2.2.25}$$

We apply this formula to the left hand side of (2.2.23) to obtain

$$\|(\sum_{k\in\mathbb{Z}} |\tilde{P}_k(f)|^2)^{1/2}\|_p^p \sim \int_{\mathbb{R}} \int_{\Gamma_{\tilde{x}}} (\sum_{k\in\mathbb{Z}} |\tilde{P}_k(f)|^2)^{p/2} ds_{\tilde{x}} d\tilde{x}. \tag{2.2.26}$$

For every fixed $\tilde{x}$, by Definition 2.5, the right hand side of (2.2.26) turns to

$$\int_{\mathbb{R}} [\int_{\mathbb{R}} (\sum_k |P_k(\tilde{f}_{\tilde{x}})(y')|^2)^{p/2} dy'] d\tilde{x}, \tag{2.2.27}$$

where $\tilde{f}_{\tilde{x}}(y') = f(g_{\tilde{x}}(y'), y')$. Then the classical Littlewood-Paley theory applies and we can bound the last expression by

$$\int_{\mathbb{R}} \|f\|_{L^p(\Gamma_{\tilde{x}})}^p d\tilde{x} \lesssim \|f\|_{L^p}^p. \tag{2.2.28}$$

For the boundedness of the adjoint operator, it suffices to prove that

$$\sum_{k\in\mathbb{Z}} \langle \tilde{P}_k^*(f), f_k \rangle \lesssim \|f\|_{L^p} \|(\sum_{k\in\mathbb{Z}} |f_k|^2)^{1/2}\|_{L^{p'}}. \tag{2.2.29}$$

First by linearity and Hölder's inequality, we derive

$$\sum_{k\in\mathbb{Z}} \langle \tilde{P}_k^*(f), f_k \rangle = \langle f, \sum_{k\in\mathbb{Z}} \tilde{P}_k(f_k) \rangle \lesssim \|f\|_{L^p} \|\sum_{k\in\mathbb{Z}} \tilde{P}_k(f_k)\|_{L^{p'}}. \tag{2.2.30}$$

Applying the co-area formula (2.2.25), we obtain

$$\|\sum_{k\in\mathbb{Z}} \tilde{P}_k(f_k)\|_{L^{p'}} \sim (\int_{\mathbb{R}} (\int_{\Gamma_{\tilde{x}}} |\sum_{k\in\mathbb{Z}} \tilde{P}_k(f_k)|^{p'} ds_{\tilde{x}}) d\tilde{x})^{1/p'}. \tag{2.2.31}$$

By the Definition 2.5, for every fixed $\tilde{x} \in \mathbb{R}$, the inner integration in the last expression turns to

$$\int_{\mathbb{R}} |\sum_{k\in\mathbb{Z}} P_k(\tilde{f}_{k,\tilde{x}})(y')|^{p'} dy', \tag{2.2.32}$$

where $\tilde{f}_{k,\tilde{x}}(y') := f_k(g_{\tilde{x}}(y'), y')$. Now the classical Littlewood-Paley theory applies and we bound the term in (2.2.32) by

$$\int_{\mathbb{R}} (\sum_{k\in\mathbb{Z}} |\tilde{f}_{k,\tilde{x}}(y')|^2)^{p'/2} dy' \lesssim \int_{\Gamma_{\tilde{x}}} (\sum_{k\in\mathbb{Z}} |f_k|^2)^{p'/2} ds_{\tilde{x}} \lesssim \|(\sum_{k\in\mathbb{Z}} |f_k|^2)^{1/2}\|^{p'}_{L^{p'}(\Gamma_{\tilde{x}})}. \tag{2.2.33}$$

Then to prove (2.2.29), we just need to integrate $d\tilde{x}$ in (2.2.33) and apply the co-area formula (2.2.25) to derive

$$\begin{aligned} \|\sum_{k\in\mathbb{Z}} \tilde{P}_k(f_k)\|_{L^{p'}} &\lesssim (\int_{\mathbb{R}} \|(\sum_{k\in\mathbb{Z}} |f_k|^2)^{1/2}\|^{p'}_{L^{p'}(\Gamma_{\tilde{x}})} d\tilde{x})^{1/p'} \\ &\lesssim \|(\sum_{k\in\mathbb{Z}} |f_k|^2)^{1/2}\|_{L^{p'}}. \end{aligned}$$

Thus we have finished the proof of Lemma 2.6.□

Now we will show how to prove the L^2 boundedness of the main term by Lemma 2.6 and Proposition 2.3: first by duality, we have

$$\begin{aligned} \|\sum_k \tilde{P}_k H_v P_k(f)\|_2 &= \sup_{\|g\|_2=1} |\langle \sum_k \tilde{P}_k H_v P_k(f), g\rangle| \\ &= \sup_{\|g\|_2=1} |\langle \sum_k H_v P_k(f), \tilde{P}_k^*(g)\rangle|. \end{aligned}$$

Applying the Cauchy-Schwartz inequality and Hölder's inequality, we can bound the last term by

$$\sup_{\|g\|_2=1} \|(\sum_k |H_v P_k(f)|^2)^{1/2}\|_2 \|(\sum_k |\tilde{P}_k^*(g)|^2)^{1/2}\|_2. \tag{2.2.34}$$

For the former term, Proposition 2.3 implies that

$$\begin{aligned} \|(\sum_k |H_v P_k(f)|^2)^{1/2}\|_2 &\leq (\sum_{k\in Z} \|H_v P_k(f)\|_2^2)^{1/2} \\ &\lesssim (\sum_{k\in\mathbb{Z}} \|P_k(f)\|_2^2)^{1/2} \lesssim \|f\|_2. \end{aligned}$$

For the latter term, Lemma 2.6 implies that

$$\|(\sum_k |\tilde{P}_k^*(g)|^2)^{1/2}\|_2 \lesssim \|g\|_2. \tag{2.2.35}$$

Thus we have proved the L^2 boundedness the main term, modulo Proposition 2.3.

As the second step, we will prove the L^2 boundedness of the commutator, which is

$$\|\sum_k (H_v P_k(f) - \tilde{P}_k H_v P_k(f))\|_2 \lesssim \|f\|_2. \tag{2.2.36}$$

To do this, we first split the operator H_v into a dyadic sum: select a Schwartz function ψ_0 such that ψ_0 is supported on $[\frac{1}{2}, \frac{5}{2}]$, let

$$\psi_l(t) := \psi_0(2^{-l}t), \tag{2.2.37}$$

by choosing ψ_0 properly, we can construct a partition of unity for $\mathbb{R}^+$, i.e.

$$\mathbb{1}_{(0,\infty)} = \sum_{l\in\mathbb{Z}} \psi_l. \tag{2.2.38}$$

Let

$$H_l h(x,y) := \int \check{\psi}_l(t) h(x-t, y-tu(P(x,y)))dt, \tag{2.2.39}$$

then the operator H_v can be decomposed into the sum

$$H_v = -\mathbb{1} + 2\sum_{l\in\mathbb{Z}} H_l. \tag{2.2.40}$$

Hence to bound the commutator, it is equivalent to bound the following

$$\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} (H_l P_k f - \tilde{P}_k H_l P_k f). \tag{2.2.41}$$

Notice that by definition, $H_l P_k f$ vanishes for $l > k$, which simplifies the last expression to

$$\sum_{l\geq 0}\sum_{k\in\mathbb{Z}} (H_{k-l} P_k f - \tilde{P}_k H_{k-l} P_k f). \tag{2.2.42}$$

By the triangle inequality, it suffices to prove

Proposition 2.7 *Under the same assumption as in Theorem 2.1, there exists $\gamma > 0$ such that*

$$\| \sum_{k \in \mathbb{Z}} (H_{k-l} P_k f - \tilde{P}_k H_{k-l} P_k f) \|_2 \lesssim 2^{-\gamma l} \|f\|_2, \tag{2.2.43}$$

with the constant independent of $l \in \mathbb{N}$.

So far, we have reduced the proof of the Main Theorem to that of Proposition 2.3 and Proposition 2.7, which we will present separately in the following sections.

2.3 Boundedness of the Lipschitz-Kakeya maximal function and proof of Proposition 2.3

Lacey and Li in their prominent work [21] have reduced the L^2 boundedness of the operator H_{v,ϵ_0} (given by (1.0.8)) to the boundedness of an operator they introduced, the so called Lipschitz-Kakeya maximal operator. As soon as this operator is bounded, we can then repeat the argument in Chapter 4 [21] to obtain Proposition 2.3 as a corollary.

Here we follow [4], where a slightly different version of the Lipschitz-Kakeya maximal operator is used, see the following Lemma 2.10. The only place in [4] where the one-variable vector field plays a special role is Lemma 6.2 in page 1037. Hence to prove Proposition 2.3, we just need to replace this lemma by Lemma 2.10, and leave the rest of the argument unchanged.

In this section we make an observation that both the boundedness of the Lipschitz-Kakeya maximal operator (Corollary 2.11) and its variant (Lemma 2.10) can be proved by adapting Bateman's argument in [3] to our case where the vector fields are constant only on Lipschitz curves.

Before defining the Lipschitz-Kakeya maximal operator, we first need to introduce several definitions.

Definition 2.8 *(popularity) For a rectangle $R \subset \mathbb{R}^2$, with $l(R)$ its length, $w(R)$ its width, we define its uncertainty interval $EX(R) \subset \mathbb{R}$ to be the interval of width $w(R)/l(R)$ and centered at slope(R). Then the popularity of the rectangle R is defined to be*

$$pop_R := |\{(x,y) \in \mathbb{R}^2 : u(P(x,y)) \in EX(R)\}|/|R|. \tag{2.3.1}$$

Definition 2.9 *Given two rectangles R_1 and R_2 in $\mathbb{R}^2$, we write $R_1 \leq R_2$ whenever $R_1 \subset CR_2$ and*

$EX(R_2) \subset EX(R_1)$, where C is some properly chosen large constant, and CR_2 is the rectangle with the same center as R_2 but dilated by the factor C.

Denote $\mathcal{R}_{\delta,\omega} := \{R \in \mathcal{R} : slope(R) \in [-1,1], pop_R \geq \delta, w(R) = \omega\}$, where $\mathcal{R}$ is the collection of all the rectangles in $\mathbb{R}^2$. Then the Lipschitz-Kakeya maximal function is defined as

$$M_{\mathcal{R}_{\delta,\omega}}(f)(x) := \sup_{x\in R\in\mathcal{R}_{\delta,\omega}} \frac{1}{|R|}\int_R |f| \tag{2.3.2}$$

Lemma 2.10 *Let u and P be the functions given in the definition of the operator H_v in* (2.2.5)*. Suppose $\mathcal{R}_0$ is a collection of pairwise incomparable (under "$\leq$") rectangles of uniform width such that for each $R \in \mathcal{R}_0$, we have*

$$\frac{|(u\circ P)^{-1}(EX(R)) \cap R|}{R} \geq \delta, (i.e.\ pop_R \geq \delta) \tag{2.3.3}$$

and

$$\frac{1}{|R|}\int_R \mathbb{1}_F \geq \lambda. \tag{2.3.4}$$

Then for each $p > 1$,

$$\sum_{R\in\mathcal{R}_0} |R| \lesssim \frac{|F|}{\delta\lambda^p}. \tag{2.3.5}$$

The same covering lemma argument as in Lemma 3.1 [3] shows the boundedness of Lacey and Li's Lipschitz-Kakeya maximal operator as a corollary of Lemma 2.10.

Corollary 2.11 *For all $p \in (1,\infty)$ we have the following bound*

$$\|M_{\mathcal{R}_{\delta,\omega}}\|_{L^p\to L^p} \leq C(p,a_0)\frac{1}{\delta} \tag{2.3.6}$$

Proof of Lemma 2.10: The proof is essentially due to Bateman [3]. Most of the argument in [3] remains, with just one minor modification in order to adapt to the family of Lipschitz curves on which the vector field is constant.

Definition 2.12 *(rectangles adapted to the vector field) For a rectangle $R \in \mathcal{R}_{\delta,\omega}$, with its two long sides lying on the parallel lines $y = kx + b_1$ and $y = kx + b_2$ for some $k \in [-1,1]$ and $b_1, b_2 \in \mathbb{R}$, define $\tilde{R}$ to be the adapted version of R, which is given by the set*

$$\{(x,y) : P(x,y) \in P(R)\} \bigcap \{(x, kx+b) : x \in \mathbb{R}, b \in [b_1,b_2]\}, \tag{2.3.7}$$

where P is the projection operator in (2.2.3)*.*

What we need to do is just to replace the rectangles R in [3] by $\tilde{R}$, and observe that the two key quantities–length and popularity of rectangles– are both preserved under the projection operator P up to a constant depending on the constant a_0 in Theorem 2.1. Hence we leave out the details and refer to [3].□

2.4 Boundedness of the commutator: Proof of Proposition 2.7

This section consists of three subsections. In the first subsection we will introduce some notations, most of which we adopt from Bateman's paper [4], with minor changes for our purpose. In the second we will use Jones' beta numbers and the Carleson embedding theorem to prove Proposition 2.7, modulo one crucial lemma which will be presented afterwards in the third subsection.

2.4.1 Time-frequency decomposition

The content of this subsection is basically taken from Bateman's paper [4], with minor changes as we are now dealing with all frequencies instead of one single frequency annulus.

Discretizing the functions: Fix $l \geq 0$, we write $\mathcal{D}_l$ as the collection of the dyadic intervals of length 2^{-l} contained in $[-2, 2]$. Fix a smooth positive function $\beta : \mathbb{R} \to \mathbb{R}$ s.t.

$$\beta(x) = 1, \forall |x| \leq 1; \beta(x) = 0, \forall |x| \geq 2. \tag{2.4.1}$$

Also choose β such that $\sqrt{\beta}$ is a smooth function. Then fix an integer c(whose exact value is unimportant), for each $\omega \in \mathcal{D}_l$, define

$$\beta_\omega(x) = \beta(2^{l+c}(x - c_{\omega_1})), \tag{2.4.2}$$

where ω_1 is the right half of ω and c_{ω_1} is its center.

Define

$$\beta_l(x) = \sum_{\omega \in \mathcal{D}_l} \beta_\omega(x), \tag{2.4.3}$$

note that

$$\beta_l(x + 2^{-l}) = \beta_l(x), \forall x \in [-2, 2 - 2^{-l}]. \tag{2.4.4}$$

Define

$$\gamma_l = \frac{1}{2}\int_{-1}^{1} \beta_l(x + t)dt, \tag{2.4.5}$$

because of the above periodicity, we know that γ_l is constant for $x \in [-1, 1]$, independent of l. Say $\gamma_l(x) = \delta > 0$, hence

$$\frac{1}{\delta}\gamma_l(x)\mathbb{1}_{[-1,1]}(x) = \mathbb{1}_{[-1,1]}(x). \tag{2.4.6}$$

Define another multiplier $\tilde{\beta} : \mathbb{R} \to \mathbb{R}$ with support in $[\frac{1}{2}, \frac{5}{2}]$ and $\tilde{\beta}(x) = 1$ for $x \in [1, 2]$. We define the corresponding multiplier on $\mathbb{R}^2$:

$$\begin{aligned}
\hat{m}_{k,\omega}(\xi, \eta) &= \tilde{\beta}(2^{-k}\eta)\beta_\omega(\frac{\xi}{\eta}) \\
\hat{m}_{k,l,t}(\xi, \eta) &= \tilde{\beta}(2^{-k}\eta)\beta_l(t + \frac{\xi}{\eta}) \\
\hat{m}_{k,l}(\xi, \eta) &= \tilde{\beta}(2^{-k}\eta)\gamma_l(\frac{\xi}{\eta})
\end{aligned}$$

Then what we need to bound can be written as

$$\begin{aligned}
\|\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} H_l P_k(f)\|_p &= \|\int_{-1}^{1}\sum_{k\in\mathbb{Z}}\sum_{l\geq 0} H_{k-l}(\frac{1}{\delta}m_{k,l} * f)dt\|_p \\
&\leq \int_{-1}^{1}\|\sum_{k\in\mathbb{Z}}\sum_{l\geq 0} H_{k-l}(\frac{1}{\delta}m_{k,l,t} * f)\|_p dt,
\end{aligned}$$

where the terms $H_l P_k$ for $l > k$ in the sum vanish as explained before.

So it suffices to prove a uniform bound on $t \in [-1, 1]$, w.l.o.g. we will just consider the case $t = 0$, which is

$$\sum_{k\in\mathbb{Z}}\sum_{l\geq 0} H_{k-l}(m_{k,l,0} * f) = \sum_{k\in\mathbb{Z}}\sum_{l\geq 0} H_{k-l}([\tilde{\beta}(2^{-k}\eta)\beta_l(\frac{\xi}{\eta})] * f). \tag{2.4.7}$$

Constructing the tiles: For each $k \in \mathbb{Z}$ and $\omega \in \mathcal{D}_l$ with $l \geq 0$, let $\mathcal{U}_{k,\omega}$ be a partition of $\mathbb{R}^2$ by rectangles of width 2^{-k} and length 2^{-k+l}, whose long side has slope θ, where $\tan\theta = -c(\omega)$, which is the center of the interval ω. If $s \in \mathcal{U}_{k,\omega}$, we will write $\omega_s := \omega$, and $\omega_{s,1}$ to be the right half of ω, $\omega_{s,2}$ the left half.

An element of $\mathcal{U}_{k,\omega}$ for some $\omega \in \mathcal{D}_l$ is called a "tile". Define $\varphi_{k,\omega}$ such that

$$|\hat{\varphi}_{k,\omega}|^2 = \hat{m}_{k,\omega}, \tag{2.4.8}$$

then $\varphi_{k,\omega}$ is smooth by our assumption on β mentioned above.

For a tile $s \in \mathcal{U}_{k,\omega}$, define

$$\varphi_s(p) := \sqrt{|s|}\varphi_{k,\omega}(p - c(s)), \tag{2.4.9}$$

where $c(s)$ is the center of s. Notice that

$$\|\varphi_s\|_2^2 = \int_{\mathbb{R}^2} |s|\varphi_{k,\omega}^2 = |s| \int_{\mathbb{R}^2} \hat{m}_{k,\omega} = 1, \tag{2.4.10}$$

i.e. φ_s is L^2 normalised.

The constructing of the tiles above by uncertainty principle is to localise the function further in space, which is realised through

Lemma 2.13 *(Page 1030 [4]) Under the above notations, for the frequency localised function $f * m_{k,\omega}$, we have the following representation:*

$$f * m_{k,\omega}(x) = \lim_{N\to\infty} \frac{1}{4N^2} \int_{[-N,N]^2} \sum_{s\in\mathcal{U}_{k,\omega}} \langle f, \varphi_s(p + \cdot)\rangle \varphi_s(p + x)dp \tag{2.4.11}$$

The above lemma allows us to pass to the model sum

$$\sum_{k\in\mathbb{Z}} \sum_{l\geq 0} H_{k-l}(f * m_{k,l,0}) = \sum_{k\in\mathbb{Z}} \sum_{l\geq 0} \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} \langle f, \varphi_s\rangle H_{k-l}(\varphi_s),$$

define

$$\psi_s = \psi_{-\log(length(s))}, \tag{2.4.12}$$

and

$$\phi_s(x, y) := \int \check{\psi}_s(t)\varphi_s(x - t, y - tu(P(x, y)))dt, \tag{2.4.13}$$

then the model sum turns to

$$\sum_{k\in\mathbb{Z}} \sum_{l\geq 0} \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} \langle f, \varphi_s\rangle \phi_s \tag{2.4.14}$$

Lemma 2.14 *We have that $\phi_s(x, y) = 0$ unless $-u(P(x, y)) \in \omega_{s,2}$.*

The proof of Lemma 2.14 is by the Plancherel theorem, we just need to observe that the frequency support of ψ_s and $\hat{\varphi}_s$ will be disjoint at the point (x, y) unless $-u(P(x, y)) \in \omega_{s,2}$.

2.4.2 Proof of Proposition 2.7

This subsection is devoted to the proof of Proposition 2.7, which is largely motivated by the proof of the $T(b)$ theorem and the boundedness of the paraproduct, see [1] and [10] for example.

In our case, unlike Bateman and Thiele's proof for the one-variable vector fields, it's no longer true that $H_v P_k f$ still has frequency in the k-th annulus. In order to get enough orthogonality for the term $H_v P_k f$ to apply the Littlewood-Paley theory, we need to subtract the term $H_v P_k f - \tilde{P}_k H_v P_k f$, which should be viewed as a family of paraproducts.

We proceed with the details of the proof. If we expand the summation on the left hand side of Proposition 2.7 with (2.4.14), what we need to bound can be rewritten as

$$\|\sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} \langle f, \varphi_s\rangle(\phi_s - \tilde{P}_k\phi_s)\|_2 \lesssim 2^{-\gamma l}\|f\|_2. \tag{2.4.15}$$

In order to use the orthogonality of different wave packets, we will prove the L^2 bound for the dual operator, which is

$$\sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} \langle h, \phi_s - \tilde{P}_k\phi_s\rangle\varphi_s. \tag{2.4.16}$$

Notice that for $s_1 \in \mathcal{U}_{k_1,\omega_1}$ and $s_2 \in \mathcal{U}_{k_2,\omega_2}$ with $(k_1, \omega_1) \neq (k_1, \omega_2)$, we have

$$\langle\varphi_{s_1}, \varphi_{s_2}\rangle = 0 \tag{2.4.17}$$

by the definition of the wavelet function φ_s in (2.4.9). Also if we know that s_1, s_2 are in the same $\mathcal{U}_{k,\omega}$, for some k and ω, then we can find $m_0, n_0 \in \mathbb{Z}$ s.t.

$$c(s_2) = c(s_1) + (m_0 \cdot l(s_1), n_0 \cdot w(s_1)) \tag{2.4.18}$$

where $c(s)$ is the center of the tile s, $l(s)$ its length and $w(s)$ its width. Then by the non-stationary phase method we know for any $N \in \mathbb{N}$, there exists a constant C_N depending only on N s.t.

$$|\langle\varphi_{s_1}, \varphi_{s_2}\rangle| \leq \frac{C_N}{(|m_0| + |n_0| + 1)^N}. \tag{2.4.19}$$

Here we want to make a remark that the exact value of N is not important, it just denotes some large number which might vary from line to line if we use the same notation later.

Applying the above two estimates (2.4.17) (2.4.19), we obtain

$$\begin{aligned}&\|\sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} \langle h, \phi_s - \tilde{P}_k\phi_s\rangle \varphi_s\|_2^2\\ &= \sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s_1\in\mathcal{U}_{k,\omega}} \sum_{s_2\in\mathcal{U}_{k,\omega}} \langle h, \phi_{s_1} - \tilde{P}_k\phi_{s_1}\rangle \langle \varphi_{s_1}, \varphi_{s_2}\rangle \langle h, \phi_{s_2} - \tilde{P}_k\phi_{s_2}\rangle.\end{aligned}$$

As we know for any $s_1, s_2 \in \mathcal{U}_{k,\omega}$ there exists $m_0, n_0 \in \mathbb{Z}$ s.t.

$$c(s_2) = c(s_1) + (m_0 \cdot l(s_1), n_0 \cdot w(s_1)), \tag{2.4.20}$$

the above sum can be rewritten as

$$\sum_{m_0,n_0\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \sum_{\omega\in\mathcal{D}_l} \sum_{s_1\in\mathcal{U}_{k,\omega}} \langle h, \phi_{s_1} - \tilde{P}_k\phi_{s_1}\rangle \langle \varphi_{s_1}, \varphi_{s_2}\rangle \langle h, \phi_{s_2} - \tilde{P}_k\phi_{s_2}\rangle \tag{2.4.21}$$

with s_1, s_2 satisfying the relation (2.4.20).

Now fix $m_0, n_0 \in \mathbb{Z}$, by the estimate in (2.4.19), we know that

$$\begin{aligned}&\sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s_1\in\mathcal{U}_{k,\omega}} |\langle h, \phi_{s_1} - \tilde{P}_k\phi_{s_1}\rangle \langle \varphi_{s_1}, \varphi_{s_2}\rangle \langle h, \phi_{s_2} - \tilde{P}_k\phi_{s_2}\rangle|\\ &\lesssim \frac{1}{(|m_0| + |n_0| + 1)^N} \sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s_1\in\mathcal{U}_{k,\omega}} |\langle h, \phi_{s_1} - \tilde{P}_k\phi_{s_1}\rangle \langle h, \phi_{s_2} - \tilde{P}_k\phi_{s_2}\rangle|,\end{aligned}$$

by the Cauchy-Schwartz inequality, the last term is bounded by

$$\frac{1}{(|m_0| + |n_0| + 1)^N} \sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle h, \phi_s - \tilde{P}_k\phi_s\rangle|^2, \tag{2.4.22}$$

then it suffices to prove that

$$\sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} \langle h, \phi_s - \tilde{P}_k\phi_s\rangle^2 \lesssim 2^{-\gamma l}\|h\|_2^2. \tag{2.4.23}$$

First to estimate every single term $\langle h, \phi_s - \tilde{P}_k\phi_s\rangle$ for a fixed tile s: denote $s_{m,n}$ to be the shift of s by (m, n) units, i.e.

$$s_{m,n} := \{(x, y) \in \mathbb{R}^2 : (x - m \cdot l(s), y - n \cdot w(s)) \in s\}, \tag{2.4.24}$$

then by the triangle inequality we know that

$$|\langle h, \phi_s - \tilde{P}_k \phi_s \rangle| \leq \sum_{m,n \in \mathbb{Z}} | \int_{s_{m,n}} h \cdot (\phi_s - \tilde{P}_k \phi_s) dy dx|. \tag{2.4.25}$$

Recall that in Definition 2.12 we use $\tilde{R}$ to denote the adapted version of the rectangle R to the family of Lipschitz curves, then clearly $\tilde{s}_{m,n} \supset s_{m,n}$. Thus

$$|\langle h, \phi_s - \tilde{P}_k \phi_s \rangle| \leq \sum_{m,n \in \mathbb{Z}} | \int_{\tilde{s}_{m,n}} h \cdot (\phi_s - \tilde{P}_k \phi_s) dy dx|. \tag{2.4.26}$$

By the co-area formula (2.2.25), we obtain

$$\begin{aligned} |\langle h, \phi_s - \tilde{P}_k \phi_s \rangle| &\leq \sum_{m,n \in \mathbb{Z}} | \int_{\tilde{s}_{m,n}} h \cdot (\phi_s - \tilde{P}_k \phi_s) dy dx| \\ &\lesssim \sum_{m,n \in \mathbb{Z}} \int_{P(s_{m,n})} \int_{\Gamma_x \cap \tilde{s}_{m,n}} |h \cdot (\phi_s - \tilde{P}_k \phi_s)| ds_x dx, \end{aligned}$$

where ds_x stands for the arc length measure of the Lipschitz curve Γ_x.

Now for the inner integration along the curve Γ_x, we do the same change of coordinates and the same parametrisation of Γ_x as in Definition 2.5, i.e. we choose the coordinates s.t. the horizontal axis is parallel to $(1, u(x))$, and represent the curve Γ_x by the Lipschitz function $g_x(\cdot)$. If we let $J(x, s_{m,n})$ denote the projection of $\Gamma_x \cap \tilde{s}_{m,n}$ on the new vertical axis, the last expression becomes

$$\sum_{m,n \in \mathbb{Z}} \int_{P(s_{m,n})} \int_{J(x, s_{m,n})} |h(g_x(y), y)(\phi_s(g_x(y), y) - P_k[\phi_s(g_x(y), y)])| dy dx. \tag{2.4.27}$$

To bound the above term, Jones' beta number will play a crucial role.

Definition 2.15 *([16]) For a Lipschitz function $A : \mathbb{R} \to \mathbb{R}$, we first take the Calderón decomposition of $a(x) = A'(x)$, which yields the representation*

$$a(x) = \sum_{I \ dyadic} a_I \psi_I(x), \tag{2.4.28}$$

where ψ_I is some mean zero function supported on $3I$, $|\psi_I'(x)| \leq |I|^{-1}$. For each dyadic interval I, let

$$\alpha_I = \sum_{|J| \geq |I|} a_I \psi_J(c_I), \tag{2.4.29}$$

where c_I stands for the center of I, denote the "average slope" of the Lipschitz curve near I, and *define the beta number*

$$\beta_0(I) := \sup_{x\in 3I} \frac{|A(x) - A(c_I) - \alpha_I(x - c_I)|}{|I|}, \tag{2.4.30}$$

and the j_0-th beta number

$$\beta_{j_0}(I) := \sup_{x\in 3j_0 I} \frac{|A(x) - A(c_I) - \alpha_I(x - c_I)|}{|I|}. \tag{2.4.31}$$

For beta numbers, we have the following Carleson condition.

Lemma 2.16 *([16]) For any Lipschitz function A, we have*

$$\sup_J \frac{1}{|J|} \sum_{I\subset J} \beta_0^2(I)|I| \lesssim \|A\|_{Lip}^2, \tag{2.4.32}$$

and also for any $j_0 \in \mathbb{N}$

$$\sup_J \frac{1}{|J|} \sum_{I\subset J} \beta_{j_0}^2(I)|I| \lesssim j_0^3 \|A\|_{Lip}^2. \tag{2.4.33}$$

After introducing Jones' beta number, we are ready to state

Lemma 2.17 *for $x \in P(s_{m,n})$, we have the following estimate:*

$$\begin{aligned}
&\int_{J(x,s_{m,n})} |h(g_x(y), y)(\phi_s(g_x(y), y) - P_k[\phi_s(g_x(y), y)])| dy \\
&\lesssim \sum_{j_0\in\mathbb{N}} \frac{2^{-3l/2}}{(|j_0| + |m| + |n| + 1)^N} \beta_{j_0}(x, s_{m,n})[h]_{x,s_{m,n}} \mathbb{1}_{\{-u(x)\in\omega_{s,2}\}}(x)
\end{aligned}$$

where $\beta_{j_0}(x, s_{m,n})$ is the j_0-th beta number for the Lipschitz curve $g_x(\cdot)$ on the interval $J(x, s_{m,n})$, $[h]_{x,s_{m,n}}$ is the average of the function h on the interval $J(x, s_{m,n})$, i.e.

$$[h]_{x,s_{m,n}} := \frac{1}{w(s)} \int_{J(x,s_{m,n})} |h(g_x(y), y)| dy. \tag{2.4.34}$$

The proof of Lemma 2.17 will be postponed to the next subsection. Substitute the estimate in Lemma 2.17 into the estimate for the term $\langle h, \phi_s - \tilde{P}_k\phi_s\rangle$, we then have that

$$\begin{aligned}
&|\langle h, \phi_s - \tilde{P}_k\phi_s\rangle| \\
&\lesssim \sum_{m,n} \int_{P(s_{m,n})} \int_{J(x,s_{m,n})} |h(g_x(y), y)(\phi_s(g_x(y), y) - P_k[\phi_s(g_x(y), y)])| dy dx \\
&\lesssim \sum_{m,n} \int_{P(s_{m,n})} \sum_{j_0\in\mathbb{N}} \frac{2^{-3l/2}}{(|j_0| + |m| + |n| + 1)^N} \beta_{j_0}(x, s_{m,n})[h]_{x,s_{m,n}} \mathbb{1}_{\{-u(x)\in\omega_{s,2}\}}(x) dx
\end{aligned}$$

hence

$$\begin{aligned}
&\sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle h, \phi_s - \tilde{P}_k\phi_s\rangle|^2 \\
&\lesssim \sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} \sum_{m,n,j_0} \frac{2^{-3l}}{(|j_0|+|m|+|n|+1)^N} \cdots \\
&\qquad\qquad \cdots \Big|\int_{P(s_{m,n})} \beta_{j_0}(x, s_{m,n})[h]_{x,s_{m,n}} \mathbb{1}_{\{-u(x)\in\omega_{s,2}\}}(x)dx\Big|^2 \\
&\lesssim \sum_{m,n,j_0} \frac{2^{-2l}}{(|j_0|+|m|+|n|+1)^N} \cdots \\
&\quad \cdots \sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} w(s) \int_{P(s_{m,n})} \beta_{j_0}^2(x, s_{m,n})[h]^2_{x,s_{m,n}} \mathbb{1}_{\{-u(x)\in\omega_{s,2}\}}(x)dx
\end{aligned}$$

Lemma 2.18 *for any fixed x, fixed m, n, j_0,*

$$\sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} w(s)\mathbb{1}_{P(s_{m,n})}(x)\beta_{j_0}^2(x, s_{m,n})[h]^2_{x,s_{m,n}} \mathbb{1}_{\{-u(x)\in\omega_{s,2}\}}(x) \lesssim j_0^3\|h\|^2_{L^2(\Gamma_x)} \tag{2.4.35}$$

Proof of Lemma 2.18: this lemma is akin to the Carleson embedding theorem, as we have the following Carleson type condition

$$\sup_{s_{m,n}} \frac{1}{|J(x, s_{m,n})|} \sum_{s'_{m,n}:J(x,s'_{m,n})\subset J(x,s_{m,n})} \beta_{j_0}^2(J(x, s'_{m,n}))w(s'_{m,n}) \lesssim j_0^3 Lip^2(\Gamma_x), \tag{2.4.36}$$

where the term $\mathbb{1}_{\{-u(x)\in\omega_{s,2}\}}$ plays such a role that, originally there are 2^l groups of dyadic rectangles

$$\bigcup_k \bigcup_{\omega\in\mathcal{D}_l} \bigcup_{s\in\mathcal{U}_{k,\omega}} \{s_{m,n}\} \tag{2.4.37}$$

in the summation $\sum_k \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}}$, which means that there are also 2^l groups of dyadic intervals

$$\bigcup_k \bigcup_{\omega\in\mathcal{D}_l} \bigcup_{s\in\mathcal{U}_{k,\omega}} \{J(x, s_{m,n})\} \tag{2.4.38}$$

which are the projections of the intersection of the dyadic rectangles with Γ_x on the vertical axis, the term $\mathbb{1}_{\{-u(x)\in\omega_{s,2}\}}$ just guarantees that there is just one such collection which has contribution, i.e. which has the right orientation in the sense of Lemma 2.14.

Then the desired estimate will just follow from the Carleson embedding theorem, which we refer to Lemma 5.1 in [1].□

Continue the calculation before the above lemma:

$$\sum_k \sum_{\omega \in \mathcal{D}_l} \sum_{s \in \mathcal{U}_{k,\omega}} |\langle h, \phi_s - \tilde{P}_k \phi_s \rangle|^2$$
$$\lesssim \sum_{m,n,j_0} \frac{2^{-2l} j_0^3}{(|j_0| + |m| + |n| + 1)^N} \int_{\mathbb{R}} \|h\|_{L^2(\Gamma_x)}^2 dx \lesssim 2^{-2l} \|h\|_2^2.$$

This finishes the proof for (2.4.23) and then Proposition 2.7 modulo Lemma 2.17, which we will present in the following subsection.

2.4.3 Proof of Lemma 2.17

We assume that $-u(x) \in \omega_{s,2}$, which means the vector $(1, u(x))$ is roughly parallel to the long side of $s_{m,n}$, otherwise the left hand side in Lemma 2.17 will also vanish due to Lemma 2.14. After the change of variables in (2.4.27), the vector $(1, u(x))$ turns to $(1, 0)$.

Proof by ignoring the tails: In order to explain how Jones' β-number appears, we first sketch the proof by ignoring the tails of the wavelet functions and the tail of the kernel of the Littlewood-Paley projection operator P_k.

By the above simplification, we only need to consider the case $m = n = 0$. What we need to "prove" becomes

$$\int_{J(x,s)} |h(g_x(y), y)(\phi_s(g_x(y), y) - P_k[\phi_s(g_x(y), y)])| dy \lesssim 2^{-3l/2} \beta_0(J(x, s))[h]_{x,s}. \tag{2.4.39}$$

For fixed x, we denote by $\tau_{x,s} y + b$ the line of "average slope" we picked in the definition of the beta number for the Lipschitz curve $g_x(\cdot)$ on the interval $J(x, s)$, for the sake of simplicity we assume $b = 0$. Moreover, as both x and s are fixed, we will also just write τ instead of $\tau_{x,s}$. Then we make the crucial observation that

$$P_k[\phi_s^x(\tau y, y)] = \phi_s^x(\tau y, y), \tag{2.4.40}$$

where

$$\phi_s^x(\tau y, y) := \int_{\mathbb{R}} \check{\psi}_s(t) \varphi_s(\tau y - t, y) dt, \tag{2.4.41}$$

due to the fact that for any function φ_s with frequency supported on the k-th annulus, if we restrict the function to a straight line, it will still have frequency supported on the k-th annulus (with one dimension less).

In comparison with the definition of ϕ_s in (2.4.13), $\phi_s^x(\tau y, y)$ is defined as the Hilbert transform along the vector $(1, u(x))$ (which is (1, 0) after the change of the variables we made in Lemma 2.4 and in the expression (2.4.27)) instead of the direction of the vector field v at the point $(\tau y, y)$.

Hence from the identity in (2.4.40) we obtain

$$\begin{aligned}
&\phi_s(g_x(y), y) - P_k[\phi_s(g_x(y), y)] \\
&= \phi_s(g_x(y), y) - P_k[\phi_s(g_x(y), y) - \phi_s^x(\tau y, y) + \phi_s^x(\tau y, y)] \\
&= \phi_s(g_x(y), y) - \phi_s^x(\tau y, y) - P_k[\phi_s(g_x(y), y) - \phi_s^x(\tau y, y)].
\end{aligned} \tag{2.4.42}$$

As we have also ignored the tails of the kernel of P_k, it is easy to see that the former and the latter terms in the last expression can essentially be handled in the same way. Hence in the following we will only consider the former term, which corresponds to the term

$$\int_{J(x,s)} |h(g_x(y), y)(\phi_s(g_x(y), y) - \phi_s^x(\tau y, y))| dy. \tag{2.4.43}$$

By the definitions of ϕ_s and ϕ_s^x, we have

$$\begin{aligned}
&|\phi_s(g_x(y), y) - \phi_s^x(\tau y, y)| \\
&= |\int_{\mathbb{R}} \check{\psi}_{k-l}(t)\varphi_s(g_x(y) - t, y)dt - \int_{\mathbb{R}} \check{\psi}_{k-l}(t)\varphi_s(\tau y - t, y)dt| \\
&= 2^{k-l}|\int_{\mathbb{R}} \check{\psi}_0(2^{k-l}t)\varphi_s(g_x(y) - t, z)dt - \int_{\mathbb{R}} \check{\psi}_0(2^{k-l}t)\varphi_s(\tau y - t, y)dt| \\
&= 2^{k-l}|\int_{\mathbb{R}} [\check{\psi}_0(2^{k-l}(t + g_x(y) - \tau y)) - \check{\psi}_0(2^{k-l}t)]\varphi_s(\tau y - t, z)dt|.
\end{aligned} \tag{2.4.44}$$

By the definition of the beta numbers, we have that

$$|g_x(y) - \tau y| \lesssim \beta_0(x, s)2^{-k}, \tag{2.4.45}$$

which implies that

$$|\check{\psi}_0(2^{k-l}(t + g_x(y) - \tau y)) - \check{\psi}_0(2^{k-l}t)| \lesssim 2^{-l}\beta_0(x, s) \tag{2.4.46}$$

by the fundamental theorem of calculus. In the end, by substituting the above estimate into (2.4.44) and (2.4.43) we obtain the desired estimate (2.4.39).

The full proof: The main idea is still the same, and the difference is that we need to be more careful with the tails of the wavelet functions and the kernel of P_k.

For fixed x, fixed m and n, denote $\tau(x, s_{m,n})y + b$ as the line of "average slope" for the Lipschitz curve $g_x(\cdot)$ on the interval $J(x, s_{m,n})$, for the sake of simplicity we assume $b = 0$. Then the crucial observation (2.4.40) becomes

$$P_k[\phi_s^x(\tau(x, s_{m,n})y, y)] = \phi_s^x(\tau(x, s_{m,n})y, y). \tag{2.4.47}$$

Hence similar to (2.4.42), we obtain from (2.4.47) that

$$\begin{aligned}
&\phi_s(g_x(y), y) - P_k[\phi_s(g_x(y), y)] \\
&= \phi_s(g_x(y), y) - \phi_s^x(\tau(x, s_{m,n})y, y) - P_k[\phi_s(g_x(y), y) - \phi_s^x(\tau(x, s_{m,n})y, y)].
\end{aligned}$$

Denote

$$I_{s_{m,n}} = |\int_{J(x,s_{m,n})} h(g_x(y), y) \cdot (\phi_s(g_x(y), y) - \phi_s^x(\tau(x, s_{m,n})y, y))dy| \tag{2.4.48}$$

and also

$$II_{s_{m,n}} = |\int_{J(x,s_{m,n})} h(g_x(y), y) \cdot P_k[\phi_s(g_x(y), y) - \phi_s^x(\tau(x, s_{m,n})y, y)]dy|. \tag{2.4.49}$$

Lemma 2.19 *Under the above notations, for* $z \in J(x, s_{m,n}) + j_0 2^{-k}$ *with* $j_0 \in \mathbb{Z}$, *we have the pointwise estimate*

$$|\phi_s(g_x(z), z) - \phi_s^x(\tau(x, s_{m,n})z, z)| \lesssim \frac{\beta_{|j_0|}(x, s_{m,n})2^k 2^{-3l/2}}{(\min\{|m| + |n|, |m| + |n| - |j_0|\} + 1)^N}. \tag{2.4.50}$$

Let us first continue the proof of Lemma 2.17: for the first term $I_{s_{m,n}}$, we take j_0 in Lemma 2.19 to be zero, then

$$|\phi_s(g_x(z), z) - \phi_s^x(\tau(x, s_{m,n})z, z)| \lesssim \frac{\beta_0(x, s_{m,n})2^k 2^{-3l/2}}{(|m| + |n| + 1)^N}, \tag{2.4.51}$$

which implies that

$$I_{s_{m,n}} \lesssim \frac{2^{-3l/2}}{(|m| + |n| + 1)^N} \beta_0(x, s_{m,n})[h]_{x,s_{m,n}}. \tag{2.4.52}$$

For the second term $II_{s_{m,n}}$, by the definition of P_k,

$$\begin{aligned}
&|P_k[\phi_s(g_x(y),y) - \phi_s^x(\tau(x,s_{m,n})y,y)]| \\
&= |\int_{\mathbb{R}} (\phi_s(g_x(z),z) - \phi_s^x(\tau(x,s_{m,n})z,z))2^k\check{\psi}_0(2^k(y-z))dz| \\
&\le |\sum_{j_0\in\mathbb{Z}} \int_{J(x,s_{m,n})+j_0 2^{-k}} (\phi_s(g_x(z),z) - \phi_s^x(\tau(x,s_{m,n})z,z))2^k\check{\psi}_0(2^k(y-z))dz|.
\end{aligned}$$

For $y \in J(x,s_{m,n})$ and $z \in J(x,s_{m,n}) + j_0 2^{-k}$, by the non-stationary phase method, we have that

$$|\check{\psi}_0(2^k(y-z))| \lesssim \frac{1}{(j_0+1)^N}, \tag{2.4.53}$$

together with the estimate in Lemma 2.19, we arrive at

$$\begin{aligned}
&|P_k[\phi_s(g_x(y),y) - \phi_s^x(\tau(x,s_{m,n})y,y)]| \\
&\lesssim \sum_{j_0\in\mathbb{Z}} \frac{\beta_{|j_0|}(x,s_{m,n})2^k 2^{-3l/2}}{(\min\{|m|+|n|, |m|+|n|-|j_0|\}+1)^N} \frac{1}{(j_0+1)^N} \\
&\lesssim \sum_{j_0\in\mathbb{Z}} \frac{\beta_{|j_0|}(x,s_{m,n})2^k 2^{-3l/2}}{(|m|+|n|+|j_0|+1)^N}.
\end{aligned}$$

Substitute the last expression into the estimate for $II_{s_{m,n}}$, we get the desired estimate. So far we have finished the proof of Lemma 2.17 except the Lemma 2.19, which we will do now.

Proof of Lemma 2.19: As x and $s_{m,n}$ are fixed now, later for simplicity we will just write τ instead of $\tau_{x,s_{m,n}}$. Notice that in the new coordinate we chose for Γ_x, the vector field along Γ_x points in the direction of $(1,0)$. Then by the definition of ϕ_s and ϕ_s^x, we have

$$\begin{aligned}
&|\phi_s(g_x(z),z) - \phi_s^x(\tau z,z)| \\
&= 2^{k-l}|\int_{\mathbb{R}} [\check{\psi}_0(2^{k-l}(t+g_x(z)-\tau z)) - \check{\psi}_0(2^{k-l}t)]\varphi_s(\tau z - t, z)dt|.
\end{aligned}$$

By the definition of the beta numbers, we have that

$$|g_x(z) - \tau z| \lesssim \beta_{|j_0|}(x,s_{m,n})2^{-k}, \tag{2.4.54}$$

which implies that

$$|\check{\psi}_0(2^{k-l}(t+g_x(z)-\tau z))-\check{\psi}_0(2^{k-l}t)|\lesssim 2^{-l}\beta_{|j_0|}(x,s_{m,n}) \tag{2.4.55}$$

by the fundamental theorem of calculus. In the end, non-stationary phase method leads to the final estimate:

$$\begin{aligned}&2^{k-l}|\int_{\mathbb{R}}[\check{\psi}_0(2^{k-l}(t+g_x(z)-\tau z))-\check{\psi}_0(2^{k-l}t)]\varphi_s(\tau z-t,z)dt|\\ &\lesssim \frac{2^{-l}\beta_{|j_0|}(x,s_{m,n})2^{\frac{k}{2}}2^{\frac{k-l}{2}}}{(\min\{|m|+|n|,|m|+|n|-|j_0|\}+1)^N}.\end{aligned}$$

Thus we have finished the proof of Lemma 2.19 and hence Lemma 2.17.

Chapter 3

Hilbert transform along measurable vector fields constant on Lipschitz curves: L^p boundedness

In this chapter, we will present the proof of Theorem 1.3 for the case $p > 3/2$. Recall that in the last chapter, we have reduced the L^2 estimate in Theorem 1.3 to the estimate (2.0.9) in Theorem 2.1. To prove the L^p estimate in Theorem 1.3, we could do the same reduction. However, here we will formulate Theorem 2.1 in a slightly different (but equivalent) way, such that it is more consistent with the language we will be using to carry out its proof. This language is the so-called δ-calculus, which has been used intensively in the Fourier restriction estimates, see [18], [13] and [8] for example. There are significant advantages of using δ-calculus, which will be explained after stating the main theorem (Theorem 3.1 below) of this chapter.

Theorem 3.1 *For vector fields $v : \mathbb{R}^2 \to \mathbb{R}^2$ of the form $(1, u(h))$ where $h : \mathbb{R}^2 \to \mathbb{R}$ is a Lipschitz function such that*

$$\|\nabla h - (1, 0)\|_\infty \le \epsilon_0 \ll 1, \tag{3.0.1}$$

and $u : \mathbb{R} \to \mathbb{R}$ is a measurable function such that

$$\|u\|_\infty \le 1, \tag{3.0.2}$$

the associated Hilbert transform, which is defined as

$$H_v f(x) := \int_{\mathbb{R}} f(x - tv(x))dt/t, \tag{3.0.3}$$

is bounded on L^p for all $p > 3/2$.

Clearly the above result is a Lipschitz perturbation of the result Theorem 1.1 by Bateman and Thiele [5]: if we take $h(x_1, x_2) = x_1$, then the vector field becomes $(1, u(h(x_1, x_2))) = (1, u(x_1))$, which is a one-variable vector field. However, we have one more assumption that $\|u\|_\infty \leq 1$. To recover the result in Theorem 1.1, we just need to apply the following anisotropic scaling

$$x_1 \to x_1, x_2 \to \lambda x_2, \tag{3.0.4}$$

and a simple limiting argument.

Let us mention the new ingredients that will be used to extend the L^2 bounds in the last chapter. Recall that in the L^2 case, the crucial ingredients are Jones' beta numbers and the adapted L^2-Littlewood-Paley theory, which is in the spirit of the work on the Cauchy integral on Lipschitz curves (for example see [10]). The techniques used in the proof of the L^2 bounds are the Hilbert space techniques as we need to use some facts like taking L^2 norm works trivially with certain square functions. Out of this reason, only L^2 bounds are obtained.

In the L^p case for p other than 2, one novelty is that we discovered a new paraproduct, which is indeed a one-parameter family of paraproducts, with each paraproduct living on one Lipschitz level curve of the vector field v. To prove the L^p bounds for the one-parameter family of paraproducts, the difficulty is how to embed each paraproduct into two dimensions without losing orthogonality. To overcome this difficulty, we need to develop an adapted L^p-Littlewood-Paley theory, which again requires a new square function as an intermediate step. This new two dimensional square function shares some common features with the bi-parameter square function. See the following crucial Lemma 3.13 and Claim 3.18.

Another difference from the L^2 case is that we will write the proof by using δ-calculus. One significant advantage of the δ-calculus, which we will see shortly in the proof, is that it allows us to express everything in terms of the function h from Theorem 3.1, instead of going back and forth between h and its inverse as in the last chapter. For example, this can be seen by comparing the crucial definition of the adapted Littlewood-Paley operator associated to the vector fields, namely by comparing Definition 2.5 in the last chapter with Definition 3.5 in the current chapter.

Organisation of chapter: In Section 3.1 we will state the strategy of the proof for Theorem 3.1. If we denote by P_k a Littlewood-Paley operator in the second variable, the main observation in

Bateman and Thiele's proof for the one-variable vector fields is that H_v commutes with P_k. In our case, this is no longer true. To recover the orthogonality, an adapted Littlewood-Paley operator was introduced in the previous chapter (see Definition 2.5 or the following Definition 3.5), which allows to split the operator H_v into a main term and a commutator term

$$\sum_{k\in\mathbb{Z}} H_v P_k(f) = \sum_{k\in\mathbb{Z}} (H_v P_k(f) - \tilde{P}_k H_v P_k(f) + \tilde{P}_k H_v P_k(f)). \tag{3.0.5}$$

The L^p $(p > 3/2)$ bounds of the main term $\sum_{k\in\mathbb{Z}} \tilde{P}_k H_v P_k(f)$ can be proved essentially by the same argument as in Bateman and Thiele [5], with just minor modifications that we will state in Section 3.2.

The main novelty is the L^p boundedness of the commutator term

$$\sum_{k\in\mathbb{Z}} (H_v P_k(f) - \tilde{P}_k H_v P_k(f)). \tag{3.0.6}$$

To achieve this, we will use the same time-frequency decomposition as in Subsection 2.4.1 in the previous chapter, and then prove in Section 3.3 that (3.0.6) is bounded on L^p for all $p > 1$.

3.1 Strategy of the proof of Theorem 3.1

We recall that if we denote by Γ the two-ended cone which forms an angle less than $\pi/4$ with the vertical axis, then by the assumption that $|u| \leq 1$, we can w.l.o.g. assume that

$$\text{supp } \hat{f} \subset \Gamma, \tag{3.1.1}$$

as for functions f with frequency supported on $\mathbb{R}^2 \setminus \Gamma$, we have that

$$H_v f(x) = H_{(1,0)} f(x), \tag{3.1.2}$$

which is the Hilbert transform along the constant vector field $(1, 0)$. But $H_{(1,0)}$ is bounded by Fubini's theorem and the L^2 boundedness of the Hilbert transform.

The rest of the proof consists of two relatively independent steps. The first step will just be an adaption of Bateman and Thiele's argument in [5] to our case. Our key observation is that both covering lemmas used there (Lemma 7 and Lemma 8) indeed hold true in our setting, from which we can derive the following proposition as a corollary by repeating the rest of the argument in [5].

Proposition 3.2 *Under the same assumptions as in Theorem 3.1, we have the following square function estimate*

$$\left\| \left(\sum_{k \in \mathbb{Z}} (H_v P_k(f))^2 \right)^{1/2} \right\|_p \lesssim \|f\|_p, \forall p > 3/2, \tag{3.1.3}$$

where P_k is the k-th Littlewood-Paley projection operator in the vertical direction defined by (2.2.11).

For the one-variable vector fields, i.e. vector fields of the form $v(x, y) = (1, u(x))$ for some measurable function u, Bateman and Thiele in [5] used (3.1.3) and the crucial observation that

$$H_v P_k = P_k H_v \tag{3.1.4}$$

to conclude the boundedness of H_v. In our case, the identity (3.1.4) is no longer true, i.e. the orthogonality between $H_v P_k f$ for different $k \in \mathbb{Z}$ is missing.

To recover the orthogonality, an adapted Littlewood-Paley operator along the level curves of the vector field was introduced in the previous chapter. This operator is in the spirit of prior work on the Cauchy integral on Lipschitz curves, but more of a bi-parameter type as we have one-parameter family of level curves.

Here we give an equivalent definition of the operator $\tilde{P}_k$ by using the language of δ-calculus. The advantage of this new definition is, compared with the one in the previous chapter, that it does not necessitate neither the change of coordinates nor the parametrisation of the Lipschitz curves, both of which can be replaced by introducing the following auxiliary function. To do this, we need several notations: for $t \in \mathbb{R}$ we define

$$\Gamma_t := \{x \in \mathbb{R}^2 : h(x) = t\}. \tag{3.1.5}$$

Moreover, we denote by v_t the value of the vector field v, which is a constant along Γ_t.

Definition 3.3 (Auxiliary Function) *For every $t \in R$, we define a new function $h_t : \mathbb{R}^2 \to \mathbb{R}$ in such a way that, if for some $y \in \Gamma_t$, we have*

$$z - y = d \cdot v_t \tag{3.1.6}$$

for some $d \in \mathbb{R}$, then we set $h_t(z) = d$.

Remark 3.4 *It is not difficult to see that*

$$|\nabla h_t| \sim 1, \ \textit{a.e. in } \mathbb{R}^2, \tag{3.1.7}$$

where the constant is independent of $t \in \mathbb{R}$.

Definition 3.5 (Adapted Littlewood-Paley Operator) *For* $x \in \mathbb{R}^2$, *we denote* $t = h(x)$. *We then define the adapted Littlewood-Paley projection operator* $\tilde{P}_k$ *restricted on the curve* Γ_t *by*

$$\tilde{P}_k f(x) := \int_{\mathbb{R}^2} \delta(h_t(y)) f(y) \check{\psi}_k((x-y)\cdot v_t^{\perp}) dy, \tag{3.1.8}$$

where $\psi_k(\cdot)$ *is given by* (2.2.10).

Remark 3.6 *We show that the above Definition 3.5 is equivalent with the Definition 2.5 in the previous chapter. To do this, we start from the new definition* (3.1.8)*: for a fixed* $t \in \mathbb{R}$, *the two vectors* v_t *and* $v_t^{\perp}$ *form a orthogonal coordinate system of the plane. Write* $y \in \mathbb{R}^2$ *in this new system as*

$$y = y_1 v_t + y_2 v_t^{\perp}, \tag{3.1.9}$$

and for the sake of simplicity we will still use the notation $y = (y_1, y_2)$. *This changes the expression in* (3.1.8) *to*

$$\begin{aligned} &\int_{\mathbb{R}^2} f(y_1, y_2)\delta(h_t(y_1, y_2))\check{\psi}_k(x_2 - y_2) dy \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(y_1, y_2)\delta(h_t(y_1, y_2)) dy_1 \right) \check{\psi}_k(x_2 - y_2) dy_2. \end{aligned} \tag{3.1.10}$$

Hence if we use the same parametrisation as the one in Definition 2.5, which is

$$\Gamma_t = \{y_2 v_t^{\perp} + g_t(y_2) v_t | y_2 \in \mathbb{R}\}, \tag{3.1.11}$$

then by the definition of the function h_t *in Definition 3.5, which implies*

$$\int_{\mathbb{R}} \delta(h_t(x)) dx_1 = 1, \tag{3.1.12}$$

the right hand side of (3.1.10) *will equal*

$$\int_{\mathbb{R}} f(g_t(y_2), y_2)\check{\psi}_k(x_2 - y_2) dy_2, \tag{3.1.13}$$

which is exactly the one given by the Definition 2.5.

Lemma 3.7 (Adapted Littlewood-Paley Theory) *For $p \in (1, \infty)$, we have the following variants of the Littlewood-Paley theorem:*

$$\|(\sum_{k\in\mathbb{Z}} |\tilde{P}_k f|^2)^{1/2}\|_p \sim \|f\|_p, \tag{3.1.14}$$

$$\|(\sum_{k\in\mathbb{Z}} |\tilde{P}_k^* f|^2)^{1/2}\|_p \sim \|f\|_p. \tag{3.1.15}$$

Proof of Lemma 3.7: This Lemma is exactly the same as Lemma 2.6 in the previous chapter. However in the following we will provide a proof by using the language of δ-calculus. By the Fubini theorem, we obtain

$$\int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} |\tilde{P}_k f|^2\right)^{p/2} = \int_{\mathbb{R}} \int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} |\tilde{P}_k f|^2\right)^{p/2} \delta(h(x) - t) dx dt. \tag{3.1.16}$$

When integrating against dx, by doing the change of variables $h(x) - t \to h_t(x)$, we can write the right hand side of the above expression as

$$\int_{\mathbb{R}} \int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} |\tilde{P}_k f|^2\right)^{p/2} \delta(h_t(x)) \frac{|\nabla h_t(x)|}{|\nabla h(x)|} dx dt. \tag{3.1.17}$$

By the bound on ∇h_t in (3.1.7) and our assumption on ∇h in (3.0.1) that

$$|\nabla h| \sim 1, \text{ a.e. in } \mathbb{R}^2, \tag{3.1.18}$$

it suffices to show that

$$\int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} |\tilde{P}_k f|^2\right)^{p/2} \delta(h_t(x)) dx \lesssim \int_{\mathbb{R}^2} |f(x)|^p \delta(h_t(x)) dx, \tag{3.1.19}$$

with a bound being independent of $t \in \mathbb{R}$.

We substitute the definition of $\tilde{P}_k$ into the left hand side of the last expression to obtain

$$\int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} \left|\int_{\mathbb{R}^2} \delta(h_t(y)) f(y) \check{\psi}_k((x-y)\cdot v_t^{\perp}) dy\right|^2\right)^{p/2} \delta(h_t(x)) dx. \tag{3.1.20}$$

The above expression can be viewed as a two dimensional Littlewood-Paley operator with the singular measure $\delta(h_t(\cdot))$, hence heuristically it is bounded by

$$\int_{\mathbb{R}^2} |f(x)|^p \delta(h_t(x))dx. \tag{3.1.21}$$

To make the above argument rigorous, we introduce the change of variables

$$x \to x_1 v_t + x_2 v_t^{\perp}, y \to y_1 v_t + y_2 v_t^{\perp}. \tag{3.1.22}$$

For the sake of simplicity, after the change of variables, we will still write $x = (x_1, x_2)$ and $y = (y_1, y_2)$. The expression in (3.1.20) hence becomes

$$\begin{aligned}
&\int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} \left| \int_{\mathbb{R}^2} \delta(h_t(y)) f(y) \check{\psi}_k(x_2 - y_2) dy \right|^2 \right)^{p/2} \delta(h_t(x))dx \\
= &\int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} \left| \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \delta(h_t(y)) f(y) dy_1 \right) \check{\psi}_k(x_2 - y_2) dy_2 \right|^2 \right)^{p/2} \delta(h_t(x))dx.
\end{aligned} \tag{3.1.23}$$

Notice that for any $x_2 \in \mathbb{R}$, we have

$$\int_{\mathbb{R}} \delta(h_t(x))dx_1 = 1. \tag{3.1.24}$$

Hence the right hand side of the last display becomes

$$\int_{\mathbb{R}} \left(\sum_{k\in\mathbb{Z}} \left| \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \delta(h_t(y)) f(y) dy_1 \right) \check{\psi}_k(x_2 - y_2) dy_2 \right|^2 \right)^{p/2} dx_2. \tag{3.1.25}$$

It is not difficult to see that the above is just a one-dimensional Littlewood-Paley square function for the function

$$\int_{\mathbb{R}} \delta(h_t(y)) f(y) dy_1, \tag{3.1.26}$$

hence it can be bounded by

$$\int_{\mathbb{R}} \left| \int_{\mathbb{R}} \delta(h_t(y)) f(y) dy_1 \right|^p dy_2 = \int_{\mathbb{R}^2} \delta(h_t(x)) |f(x)|^p dx. \tag{3.1.27}$$

So far we have finished the proof of (3.1.19), thus (3.1.14). For the second equivalence relation (3.1.15), the proof is similar, hence we leave it out. □

To proceed, we follow the same idea as in (2.2.21) in the L^2 case and split the operator into two terms,

$$\sum_{k\in\mathbb{Z}} H_v P_k(f) = \sum_{k\in\mathbb{Z}} (H_v P_k(f) - \tilde{P}_k H_v P_k(f) + \tilde{P}_k H_v P_k(f)). \tag{3.1.28}$$

Then by the triangle inequality, we have

$$\|\sum_{k\in\mathbb{Z}} H_v P_k(f)\|_p \lesssim \|\sum_{k\in\mathbb{Z}} (H_v P_k(f) - \tilde{P}_k H_v P_k(f))\|_p + \|\sum_{k\in\mathbb{Z}} \tilde{P}_k H_v P_k(f)\|_p. \tag{3.1.29}$$

We call the second term the main term, and the first term the commutator term.

To bound the main term, we first use duality to write the L^p norm into

$$\begin{aligned} \|\sum_{k\in\mathbb{Z}} \tilde{P}_k H_v P_k(f)\|_p &= \sup_{\|g\|_{p'}=1} |\langle \sum_{k\in\mathbb{Z}} \tilde{P}_k H_v P_k(f), g\rangle| \\ &= \sup_{\|g\|_{p'}=1} |\sum_{k\in\mathbb{Z}} \langle H_v P_k(f), \tilde{P}_k^*(g)\rangle|. \end{aligned}$$

Then by Cauchy-Schwartz and Hölder's inequality, we bound the right hand side by

$$\begin{aligned} &\sup_{\|g\|_{p'}=1} \int (\sum_{k\in\mathbb{Z}} |H_v P_k(f)|^2)^{1/2} (\sum_{k\in\mathbb{Z}} |\tilde{P}_k^*(g)|^2)^{1/2} \\ \lesssim &\sup_{\|g\|_{p'}=1} \|(\sum_{k\in\mathbb{Z}} |H_v P_k(f)|^2)^{1/2}\|_p \|(\sum_{k\in\mathbb{Z}} |\tilde{P}_k^*(g)|^2)^{1/2}\|_{p'}. \end{aligned} \tag{3.1.30}$$

In the end, by applying Proposition 3.2 to the former term in the last expression and Lemma 3.7 to the latter term, we get the desired bound

$$(3.1.30) \lesssim \sup_{\|g\|_{p'}=1} \|f\|_p \|g\|_{p'} = \|f\|_p. \tag{3.1.31}$$

Now we turn to the commutator term. Before explaining the idea of estimating the commutator term, we recall some notations from the previous chapter. Select a Schwartz function ψ_0 such that ψ_0 is supported on $[\frac{1}{2}, \frac{5}{2}]$, let

$$\psi_l(t) := \psi_0(2^{-l}t). \tag{3.1.32}$$

By choosing ψ_0 properly, we can construct a partition of unity for $\mathbb{R}^+$, i.e.

$$\mathbb{1}_{(0,\infty)} = \sum_{l\in\mathbb{Z}} \psi_l. \tag{3.1.33}$$

Let

$$H_l f(x) := \int \check{\psi}_l(t) f(x - tv(x)) dt. \tag{3.1.34}$$

Then the operator H_v can be decomposed into the sum

$$H_v = -\mathbb{1} + 2\sum_{l\in\mathbb{Z}} H_l. \tag{3.1.35}$$

We continue to explain the strategy of proving the L^p boundedness of the commutator term, which is

$$\|\sum_{k\in\mathbb{Z}} (H_v P_k(f) - \tilde{P}_k H_v P_k(f))\|_p \lesssim \|f\|_p. \tag{3.1.36}$$

By the dyadic decomposition in (3.1.35), this is equivalent to bound the following

$$\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} (H_l P_k f - \tilde{P}_k H_l P_k f). \tag{3.1.37}$$

Notice that by definition, $H_l P_k f$ vanishes for $l > k$, which simplifies the last expression to

$$\sum_{l\geq 0}\sum_{k\in\mathbb{Z}} (H_{k-l} P_k f - \tilde{P}_k H_{k-l} P_k f). \tag{3.1.38}$$

So by the triangle inequality it suffices to prove

Proposition 3.8 *Under the same assumptions as in Theorem 3.1, for any $p \in (1,\infty)$, there exists a constant $\gamma_p > 0$ such that*

$$\|\sum_{k\in\mathbb{Z}} (H_{k-l} P_k(f) - \tilde{P}_k H_{k-l} P_k(f))\|_p \lesssim 2^{-\gamma_p l}\|f\|_p, \tag{3.1.39}$$

with the constant being independent of $l \in \mathbb{N}$.

The idea of proving endpoint estimates like the $L^\infty \to BMO$ estimate will probably not work as the output of the operator H_v is so rough that it is only measurable across the family of Lipschitz level curves, in another word, the orthogonality between different tiles is missing.

To recover the orthogonality at the level of the L^2 estimate, the argument in the previous chapter relies heavily on the fact that taking L^2 norm works perfectly (also trivially) with the square function. Hence we could expand certain square summation and apply Hölder's inequality to turn the problem to the analysis on every single Lipschitz curve.

However, in the L^p estimate for $p \neq 2$, this strategy does not work, and instead we will invoke a new square function as an intermediate step. This square function is similar to the square function in the product space $\mathbb{R} \times \mathbb{R}$.

Remark 3.9 *Although the endpoint $L^\infty \to BMO$ estimate might not work for* (3.1.39) *with the classical BMO space, we still hope that there would be some variants, possibly similar to the fiber-wise Hardy and BMO spaces in [6] and [19], which will act as the right substitutes for the endpoint theory.*

Remark 3.10 *For the one-variable vector fields $v(x_1, x_2) = (1, u(x_2))$, it was proved in [11], under some convexity and curvature assumptions on the function $u : \mathbb{R} \to \mathbb{R}$, that the associated Hilbert transform and maximal function map $H^1_{prod}(\mathbb{R} \times \mathbb{R})$ to L^1, where $H^1_{prod}(\mathbb{R} \times \mathbb{R})$ denotes the product Hardy space.*

However, it was also pointed out that this might not be the right endpoint theory, and some new underlying Calderon-Zygmund theory is to be expected. See Remark (*iii*) *in Page 597 in [11].*

3.2 Boundedness of the main term: Proof of Proposition 3.2

The goal of this section is to make an observation that Bateman and Thiele's square function estimate (see (2.1) in [5]) for the one-variable vector fields, which is

$$\|(\sum_{k\in\mathbb{Z}}(H_v P_k(f))^2)^{1/2}\|_p \lesssim \|f\|_p, \forall p > 3/2, \tag{3.2.1}$$

works equally well for our case, with just minor modifications. Indeed, the proof of the estimate (3.2.1) is reduced by Bateman and Thiele in [5] to three covering lemmas (Lemma 7 and Lemma 8 in [5], Lemma 6.2 in [4]), and our observation is that all these covering lemmas still hold true for the case where the vector fields are constant only on Lipschitz curves instead of vertical lines.

Before stating the covering lemmas and the modification that we will make in the proof, we first recall several notations from Definition 2.8 and Definition 2.9.

For a rectangle $R \subset \mathbb{R}^2$, with l_R its length, w_R its width, we define its uncertainty interval $EX(R) \subset \mathbb{R}$ to be the interval of width w_R/l_R and centered at slope(R). Denote by $E(R)$ the collection of the points $x \in R$ s.t. the vector $v(x) = (1, u(h(x)))$ points roughly in the same direction as the long side of R:

$$E(R) = \{x \in R : u(h(x)) \in EX(R)\}. \tag{3.2.2}$$

Then the *popularity* of the rectangle R is defined to be

$$pop_R := |E(R)|/|R|. \tag{3.2.3}$$

Here u and h are the two functions in Theorem 3.1.

Now we are ready to state the key covering lemmas:

Lemma 3.11 *(Lemma 7 in [5]) Under the same assumptions as in Theorem 3.1, let* $\delta > 0$ *and* $q > 1$*, let* $G \subset \mathbb{R}^2$ *be a measurable set and* $\mathcal{R}$ *be a finite collection of rectangles such that*

$$|E(R) \cap G| \geq \delta |G| \tag{3.2.4}$$

for each $R \in \mathcal{R}$*. Then*

$$|\bigcup_{R \in \mathcal{R}} R| \lesssim \delta^{-q}|G|. \tag{3.2.5}$$

Lemma 3.12 *(Lemma 8 in [5]) Under the same assumptions as in Theorem 3.1, let* $0 < \sigma, \delta \leq 1$*, let* H *be a measurable set, and let* $\mathcal{R}$ *be a finite collection of rectangles such that for each* $R \in \mathcal{R}$ *we have*

$$pop_R \geq \sigma, |H \cap R| \geq \delta |R|. \tag{3.2.6}$$

Then

$$|\bigcup_{R \in \mathcal{R}} R| \lesssim \sigma^{-1}\delta^{-2}|H|. \tag{3.2.7}$$

To prove these covering lemmas, similar to the proof of Lemma 2.10, one just need to replace the usual rectangles by the “rectangles” adapted to the vector fields, which are given in Definition 2.12, and run the same argument as in Bateman and Thiele in [5].

These two lemmas were used to give an upper bound on the size of the exceptional sets around which the rectangles have either large size or large density. After excluding the exceptional sets, the argument in [5], together with [4](which also works equally well for our case as has been pointed out in section 2.3), will lead to the square function estimate, i.e. Proposition 3.2.

3.3 Boundedness of the commutator term: Proof of Proposition 3.8

In this section we intend to prove that for any $p > 1$, there exists $\gamma_p > 0$ such that

$$\left\| \sum_{k\in\mathbb{Z}} \Big(H_{v,k-l} P_k(f) - \tilde{P}_k H_{v,k-l} P_k(f) \Big) \right\|_p \lesssim 2^{-\gamma_p l} \|f\|_p. \tag{3.3.1}$$

If we expand the left hand side of the last expression to a model sum by the notations in Subsection 2.4.1, (3.3.1) becomes

$$\left\| \sum_{k\in\mathbb{Z}} \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} \langle f, \varphi_s\rangle (\phi_s - \tilde{P}_k \phi_s) \right\|_p \lesssim 2^{-\gamma_p l} \|f\|_p. \tag{3.3.2}$$

Observe that for a fixed point $x \in \mathbb{R}^2$, by Lemma 2.14, the expression

$$\sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} \langle f, \varphi_s\rangle (\phi_s - \tilde{P}_k \phi_s)(x) \tag{3.3.3}$$

can be non-zero for at most one $\omega \in \mathcal{D}_l$, which implies that

$$\begin{aligned} &\left\| \sum_{k\in\mathbb{Z}} \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} \langle f, \varphi_s\rangle (\phi_s - \tilde{P}_k \phi_s) \right\|_p \\ \lesssim &\left(\sum_{\omega\in\mathcal{D}_l} \int_{\mathbb{R}^2} \left| \sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} \langle f, \varphi_s\rangle (\phi_s - \tilde{P}_k \phi_s) \right|^p \right)^{1/p} \end{aligned} \tag{3.3.4}$$

From the right hand side of the above inequality, we see that (3.3.2) is reduced to separate $\omega \in \mathcal{D}_l$. Hence we just need to do the estimate for each ω separately. To be precise, we will prove

Lemma 3.13 *Under the above notations, we have*

$$\left\| \sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} \langle f, \varphi_s\rangle (\phi_s - \tilde{P}_k \phi_s) \right\|_p \lesssim 2^{-l} \|P_\omega f\|_p, \tag{3.3.5}$$

where P_ω is the frequency projection operator given by

$$\mathcal{F} P_\omega f(\xi_1, \xi_2) = \beta_\omega\left(\frac{\xi_1}{\xi_2}\right) \mathcal{F} f(\xi_1, \xi_2), \tag{3.3.6}$$

and the constant in (3.3.5) *is independent of* $\omega \in \mathcal{D}_l$*.*

Lemma 3.14 *We have the following bounds for the multiplier* β_ω*:*

$$\|P_\omega f\|_p \lesssim \|f\|_p, \tag{3.3.7}$$

for all $p \in (1, \infty)$*, with the constant being independent of* ω*.*

Finishing the proof of Proposition 3.8: we substitute the estimates in Lemma 3.13 and Lemma 3.14 into (3.3.4) to obtain

$$\begin{aligned} &\left\| \sum_{k\in\mathbb{Z}} \sum_{\omega\in\mathcal{D}_l} \sum_{s\in\mathcal{U}_{k,\omega}} \langle f, \varphi_s\rangle (\phi_s - \tilde{P}_k \phi_s) \right\|_p \\ &\lesssim \left(\sum_{\omega\in\mathcal{D}_l} 2^{-pl} \|P_\omega f\|_p^p \right)^{1/p} \lesssim 2^{-\frac{p-1}{p}\cdot l} \|f\|_p, \end{aligned} \tag{3.3.8}$$

which finishes the proof of Proposition 3.8.□

Remark 3.15 *It has been proved by Demeter and Di Plinio in [12] that*

$$\left(\sum_{\omega\in\mathcal{D}_l} \|P_\omega f\|_p^p \right)^{1/p} \lesssim \|f\|_p, \tag{3.3.9}$$

for $p \geq 2$*, with the constant being independent of* $l \in \mathbb{N}$*. This will provide a better exponential decay in* l *in the last inequality in* (3.3.8)*. However, here we do not need such orthogonality estimate but simply a triangle inequality.*

3.3.1 Proof of Lemma 3.14

We first reduce the estimate to one single $\omega \in \mathcal{D}_l$ by applying the shearing transform (the following (3.3.10)). Suppose for the moment that we have proved (3.3.5) for $\omega = [0, 2^{-l}]$, by doing the following change of variables

$$x_1 \to x_1, x_2 \to x_2 + \lambda x_1, \tag{3.3.10}$$

for the function f, the frequency variables are transformed into

$$\xi_1 \to \xi_1 - \lambda\xi_2, \xi_2 \to \xi_2. \tag{3.3.11}$$

This linear change of variables turns

$$P_{\omega'} f(\xi_1, \xi_2) = \mathcal{F}^{-1}\left(\beta_{\omega'}(\frac{\xi_1}{\xi_2})\hat{f}(\xi_1, \xi_2)\right), \tag{3.3.12}$$

which is the term on the left hand side of (3.3.7), into

$$\mathcal{F}^{-1}\left(\beta_{\omega'}(\frac{\xi_1}{\xi_2})\hat{f}(\xi_1 - \lambda\xi_2, \xi_2)\right). \tag{3.3.13}$$

If we denote

$$\tilde{\xi_1} := \xi_1 - \lambda\xi_2, \tilde{\xi_2} := \xi_2, \tag{3.3.14}$$

the multiplier in (3.3.13) turns to

$$\beta_{\omega'}(\frac{\tilde{\xi_1} + \lambda\tilde{\xi_2}}{\tilde{\xi_2}}) = \beta(2^{l+c}\frac{\tilde{\xi_1}}{\tilde{\xi_2}} + \lambda 2^{l+c} - 2^{l+c}c_{\omega_1'}). \tag{3.3.15}$$

So far it becomes clear that by taking λ in (3.3.15) properly, we can apply the change of variables (3.3.10) to turn the projection operator $P_{\omega'} f$ for an arbitrary $\omega \in \mathcal{D}_l$ to $P_\omega f$, where $\omega = [0, 2^{-l}]$.

Next, we will reduce the estimate for all $l \in \mathbb{N}$ to the one simply for $l = 0$. This can be done by applying the following anisotropic scaling symmetry:

$$x_1 \to \lambda x_1, x_2 \to x_2, \tag{3.3.16}$$

for the function f. Under the above change of variables, the Fourier transform of f is transformed from $\hat{f}(\xi_1, \xi_2)$ to

$$\frac{1}{\lambda}\hat{f}(\frac{\xi_1}{\lambda}, \xi_2). \tag{3.3.17}$$

Correspondingly, the function $P_\omega f$ is changed to

$$\begin{aligned} &\int \beta_\omega(\frac{\xi_1}{\xi_2})\frac{1}{\lambda}\hat{f}(\frac{\xi_1}{\lambda}, \xi_2)e^{ix_1\xi_1 + ix_2\xi_2}d\xi_1 d\xi_2 \\ =& \int \beta_\omega(\frac{\lambda\xi_1}{\xi_2})\hat{f}(\xi_1, \xi_2)e^{i\lambda x_1\xi_1 + ix_2\xi_2}d\xi_1 d\xi_2. \end{aligned} \tag{3.3.18}$$

Hence the multiplier $\beta_\omega(\xi_1/\xi_2)$ has the same L^p norm with $\beta_\omega(\lambda\xi_1/\xi_2)$. However, by the definition of β_ω, we have

$$\beta_\omega(\frac{\lambda\xi_1}{\xi_2}) = \beta(\frac{2^{l+c}\lambda\xi_1}{\xi_2} - 2^{l+c}c_{\omega_1}), \tag{3.3.19}$$

which means that if we take $\lambda = 2^{-l}$, the right hand side of the last expression becomes $\beta_{\omega_0}(\xi_1/\xi_2)$ where $\omega_0 = [0, 1]$.

After the above reductions, we just need to prove (3.3.7) with $\omega_0 = [0, 1]$. For $p = 2$, the estimate is trivial due to Plancherel's theorem. For $p \neq 2$, if we denote by P_k a Littlewood-Paley projection operator in the second variable, then by the Littlewood-Paley theory, we obtain

$$\|P_{\omega_0} f\|_p \lesssim \left\| \left(\sum_k |P_k P_{\omega_0} f|^2 \right)^{1/2} \right\|_p . \tag{3.3.20}$$

By the classical Calderon-Zygmund theory, it is not difficult to prove that

$$\left\| \left(\sum_k |P_k P_{\omega_0} f|^2 \right)^{1/2} \right\|_{BMO} \lesssim \|f\|_\infty, \tag{3.3.21}$$

and

$$\left\| \left(\sum_k |P_k P_{\omega_0} f|^2 \right)^{1/2} \right\|_1 \lesssim \|f\|_{H^1}. \tag{3.3.22}$$

Hence by interpolation, we obtain the desired estimate for all $p \in (1, \infty)$. So far we have finished the proof of Lemma 3.14. $\square$

3.3.2 Proof of Lemma 3.13

By the same shearing transform as in (3.3.10), we can reduce the estimate (3.3.5) for different ω to the one for a fixed ω, say $\omega = [0, 2^{-l}]$. To prove (3.3.5), by invoking duality, it is equivalent to prove that

$$\int_{\mathbb{R}^2} \sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f, \varphi_s\rangle| \left| \left(\phi_s - \tilde{P}_k \phi_s \right) \cdot g \right| \lesssim 2^{-l} \|f\|_p, \tag{3.3.23}$$

where the function g satisfies $\|g\|_{p'} \leq 1$. By the Fubini theorem, the left hand side of (3.3.23) is equal to

$$\begin{aligned} &\int_{\mathbb{R}} \int_{\mathbb{R}^2} \sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f, \varphi_s\rangle| \left| \left(\phi_s(x) - \tilde{P}_k \phi_s(x) \right) \cdot g(x) \right| \delta(h(x) - t) dx dt \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}^2} \sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f, \varphi_s\rangle| \left| \left(\phi_s(x) - \tilde{P}_k \phi_s(x) \right) \cdot g(x) \right| \delta(h_t(x)) \frac{|\nabla h_t(x)|}{|\nabla h(x)|} dx dt. \end{aligned} \tag{3.3.24}$$

By the bound on ∇h_t in (3.1.7) and our assumption on ∇h in Theorem 3.1, the right hand side of (3.3.24) can be bounded by

$$\int_{\mathbb{R}} \int_{\mathbb{R}^2} \sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f, \varphi_s\rangle| \left| \left(\phi_s(x) - \tilde{P}_k \phi_s(x)\right) \cdot g(x) \right| \delta(h_t(x)) dx dt. \tag{3.3.25}$$

If we denote by $s_{m,n}$ the translation of the tile s by (m, n) units, which is

$$s_{m,n} := s - (m \cdot l_s, n \cdot w_s), \tag{3.3.26}$$

then the above (3.3.25) is equal to

$$\sum_{m,n} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f, \varphi_s\rangle| \mathbb{1}_{s_{m,n}}(x) \left| \left(\phi_s(x) - \tilde{P}_k \phi_s(x)\right) \cdot g(x) \right| \delta(h_t(x)) dx dt. \tag{3.3.27}$$

By the notion of the adapted rectangles in Definition 2.12, we can replace $s_{m,n}$ by the slightly enlarged "rectangle" $\tilde{s}_{m,n}$ as from the definition it is clear that $\tilde{s}_{m,n} \supset s_{m,n}$. Moreover, in the following, we will only focus on the term $m = n = 0$, as the other terms appear as the tail terms by the non-stationary phase method.

The pointwise estimate in the following Lemma 3.16 (which has been essentially contained in Lemma 2.17) will play a crucial role in the forthcoming calculation. To state this estimate, we need to make some preparations: for a fix $t \in \mathbb{R}$, we use the new coordinates system given by $(v_t, v_t^{\perp})$. For a tile s, we use $J(t, s)$ to denote the projection of $\Gamma_t \cap \tilde{s}$ on the new vertical axis $v_t^{\perp}$. Moreover for the interval $J(t, s)$, we let $J^D(t, s)$ denote one of the dyadic intervals (at most two) on the vertical axis such that

$$|J^D(t, s)| \in (8 \cdot |J(t, s)|, 16 \cdot |J(t, s)|] \tag{3.3.28}$$

and

$$|J^D(t, s) \cap J(t, s)| \geq |J(t, s)|/2. \tag{3.3.29}$$

For the dyadic interval $J^D(t, s)$, we let $\Phi_{J^D(t,s)}$ denote the associated L^2 normalised Haar function.

Lemma 3.16 *Fix $t \in \mathbb{R}$ and $s \in \mathcal{U}_{k,\omega}$ for some $\omega \in \mathcal{D}_l$, for $x \in \Gamma_t \cap \tilde{s}$, we have the pointwise estimate*

$$|\phi_s(x) - \tilde{P}_k \phi_s(x)| \lesssim \sum_{j_0\in\mathbb{N}} \frac{2^{-3l/2} 2^k \beta_{j_0}(J^D(t, s))}{(j_0 + 1)^N}, \tag{3.3.30}$$

where $\beta_{j_0}(J^D(t,s))$ denotes the j_0-th beta number of Γ_t near the dyadic interval $J^D(t,s)$.

Remark 3.17 *The proof of the above Lemma 3.16 has been covered by the proof of Lemma 2.17. However, the argument there relies on those unnecessary parameters and auxiliary functions that we want to avoid by doing δ-calculus. As we have promised in the beginning of this chapter that we will carry out the whole argument in the language of δ-calculus completely, we should also be able to prove Lemma 3.16 by doing so. This is postponed to the next subsection.*

Substitute the above estimate into the right hand side of (3.3.27) with $m = n = 0$, we obtain

$$\sum_{j_0} \frac{2^{-l}}{(j_0+1)^N} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f,\varphi_s\rangle| 2^k 2^{-l/2} \mathbb{1}_{\tilde{s}}(x) \beta_{j_0}(J^D(t,s)) \cdot |g(x)| \delta(h_t(x)) dx dt. \tag{3.3.31}$$

To proceed, we need the following

Claim 3.18 *Fix $t \in \mathbb{R}$, we have the following estimate*

$$\begin{aligned}
&\int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} 2^k 2^{-l/2} |\langle f,\varphi_s\rangle| \mathbb{1}_{\tilde{s}}(x) \beta_{j_0}(J^D(t,s)) \right) g(x)\delta(h_t(x))dx \\
&\lesssim j_0^{3/2} \left(\int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f,\varphi_s\rangle|^2 \chi_s^2(x) \right)^{p/2} \delta(h_t(x))dx \right)^{1/p} \left(\int_{\mathbb{R}_2} |g(x)|^{p'} \delta(h_t(x))dx \right)^{1/p'},
\end{aligned}$$

where for $x = (x_1, x_2)$,

$$\chi_s(x_1,x_2) := \frac{|s|^{-1/2}}{(1 + (\frac{x_1 - c_{s,1}}{l_s})^2 + (\frac{x_2 - c_{s,2}}{w_s})^2)^5}, \tag{3.3.32}$$

with $c_s = (c_{s,1}, c_{s,2})$ denoting the center of s, $l_s = 2^{-k+l}$ the length and $w_s = 2^{-k}$ the width.

We postpone the proof of the Claim 3.18 till the end of this subsection and continue with the estimate of the term (3.3.31). By Claim (3.18) and by applying Hölder's inequality to $\int_{\mathbb{R}} dt$, the expression in (3.3.31) can be bounded by

$$\begin{aligned}
&\sum_{j_0} \frac{j_0^{3/2} \cdot 2^{-l}}{< j_0 >^N} \left(\int_{\mathbb{R}} \int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f,\varphi_s\rangle|^2 \chi_s^2(x) \right)^{p/2} \delta(h_t(x))dxdt \right)^{1/p} \\
&\lesssim 2^{-l} \cdot \left\| \left(\sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f,\varphi_s\rangle|^2 \chi_s^2(x) \right)^{1/2} \right\|_p .
\end{aligned} \tag{3.3.33}$$

To bound the last expression, we need the following

Lemma 3.19 *We have the following variant of the square function estimate*

$$\left\| \left(\sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f,\varphi_s\rangle|^2 \chi_s^2(x) \right)^{1/2} \right\|_p \lesssim \|f\|_p. \tag{3.3.34}$$

Finishing the proof of Lemma 3.13: it is straightforward that, combined with (3.3.33), Lemma 3.19 finishes the estimate of the expression (3.3.31), thus the proof of Lemma 3.13. □

Proof of Lemma 3.19: recall that in the estimate (3.3.34), we have $\omega = [0, 2^{-l}]$. Now we want to reduce the estimate to the case $\omega_0 = [0,1]$ by applying the anisotropic scaling

$$x_1 \to 2^l x_1, x_2 \to x_2. \tag{3.3.35}$$

Under the above change of variables, as has been explained in the proof of Lemma 3.14, φ_s for some $s \in \mathcal{U}_{k,\omega}$ is changed to $\varphi_{s'}$ for the corresponding $s' \in \mathcal{U}_{k,\omega_0}$ with $\omega_0 = [0,1]$. Moreover, the function χ_s will also behave in the same way:

$$\begin{aligned}\chi_s(2^l x_1, x_2) &= \frac{|s|^{-1/2}}{(1+(\frac{2^l x_1 - c_{s,1}}{l_s})^2 + (\frac{x_2-c_{s,2}}{w_s})^2)^5} \\ &= \frac{|s|^{-1/2}}{(1+(\frac{x_1 - 2^{-l} c_{s,1}}{2^{-l} l_s})^2 + (\frac{x_2-c_{s,2}}{w_s})^2)^5}.\end{aligned} \tag{3.3.36}$$

Recall that $l_s = 2^l w_s$, hence the right hand side of (3.3.36) becomes a bump function with main support on a cube of side length w_s, which means that $\chi_s(2^l x_1, x_2)$ is equal to $\chi_{s'}$ for some $s' \in \mathcal{U}_{k,\omega_0}$ up to a normalisation factor.

After the above reduction, we just need to prove (3.3.34) for $\omega = [0,1]$. For the case $p = 2$, by the orthogonality of the wavelet functions, we obtain

$$\left(\int_{\mathbb{R}^2} \sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f,\varphi_s\rangle|^2 \chi_s^2 \right)^{1/2} \lesssim \|f\|_2. \tag{3.3.37}$$

Moreover, by the classical Calderon-Zygmund theory, it is not difficult to prove the following endpoint estimates

$$\left\| \left(\sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f, \varphi_s\rangle|^2 \chi_s^2 \right)^{1/2} \right\|_{BMO} \lesssim \|f\|_\infty, \tag{3.3.38}$$

and

$$\left\| \left(\sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f, \varphi_s\rangle|^2 \chi_s^2 \right)^{1/2} \right\|_{1} \lesssim \|f\|_{H^1}. \tag{3.3.39}$$

Hence by interpolation, we can obtain all the expected L^p estimate for (3.3.34) in the above Lemma 3.19. □

Proof of Claim 3.18: For a fixed $t \in \mathbb{R}$, for the summation on the left hand side of the estimate in Claim 3.18, we observe that

$$\sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} = \sum_{s: s\cap\Gamma_t \neq \emptyset}, \tag{3.3.40}$$

as the term $\mathbb{1}_{\tilde{s}}(x)$ will vanish if $s \cap \Gamma_t = \emptyset$. We use the new coordinate system $(v_t, v_t^\perp)$, and write $x = x_1 v_t + x_2 v_t^\perp$, which will still be denoted as $x = (x_1, x_2)$ for the sake of simplicity. This turns the left hand side of the estimate in Claim 3.18 into

$$\begin{aligned} &\int_{\mathbb{R}}\int_{\mathbb{R}} \left(\sum_{s: s\cap\Gamma_t\neq\emptyset} |\langle f, \varphi_s\rangle| \mathbb{1}_{\tilde{s}}(x_1,x_2) \beta_{j_0}(J^D(t,s)) 2^k 2^{-l/2} \right) g(x_1,x_2)\delta(h_t(x_1,x_2))dx_1dx_2 \\ &= \sum_{s: s\cap\Gamma_t\neq\emptyset} 2^k 2^{-l/2} |\langle f, \varphi_s\rangle| \beta_{j_0}(J^D(t,s)) \int_{\mathbb{R}}\int_{\mathbb{R}} g(x_1,x_2)\mathbb{1}_{\tilde{s}}(x_1,x_2)\delta(h_t(x_1,x_2))dx_1dx_2 \end{aligned} \tag{3.3.41}$$

Notice that the integration on the right hand side of (3.3.41) can be estimated in the following way

$$\begin{aligned} &\left| \int_{\mathbb{R}}\int_{\mathbb{R}} g(x_1,x_2)\mathbb{1}_{\tilde{s}}(x_1,x_2)\delta(h_t(x_1,x_2))dx_1dx_2 \right| \\ &\lesssim 2^{-k} \left[\int_{\mathbb{R}} g(x_1,\cdot)\delta(h_t(x_1,\cdot))dx_1 \right]_{2J^D(t,s)}, \end{aligned}$$

where for a function $G : \mathbb{R} \to \mathbb{R}$, $[G(\cdot)]_J$ denotes the average of the function G on the interval $J \subset \mathbb{R}$.

Substitute the above bound into the right hand side of (3.3.41), we obtain the following bound

$$\sum_{s: s\cap\Gamma_t\neq\emptyset} 2^{-l/2} |\langle f, \varphi_s\rangle| \beta_{j_0}(J^D(t,s)) \left[\int_{\mathbb{R}} g(x_1,\cdot)\delta(h_t(x_1,\cdot))dx_1 \right]_{2J^D(t,s)}.$$

To proceed, the idea is to view the above expression as a paraproduct. To do this, we need to find the right function such that it has the wavelet coefficient $2^{-l/2}|\langle f, \varphi_s\rangle| w_s^{-1/2}$, where $w_s = 2^{-k}$

denotes the width of the tile s. This can be achieved by defining a function $F_t : \mathbb{R} \to \mathbb{R}$ such that

$$F_t(x_2) = \sum_{s:s\cap\Gamma_t\neq\emptyset} 2^{-l/2} w_s^{-1/2} \langle f, \varphi_s\rangle \Phi_{J^D(t,s)}(x_2), \tag{3.3.42}$$

where $\Phi_{J^D(t,s)}$ denotes the L^2 normalised Haar function associated to the dyadic interval $J^D(t,s)$.

By the L^p boundedness of the paraproduct (see [1] for example) and Jones' beta number condition that

$$\sup_s \frac{1}{|J^D(t,s)|} \sum_{s':J^D(t,s')\subset J^D(t,s)} \beta_{j_0}^2(J^D(t,s'))w_s \lesssim j_0^3, \tag{3.3.43}$$

we obtain for any fixed $t \in \mathbb{R}$ that

$$\begin{aligned}
&\sum_{s:s\cap\Gamma_t\neq\emptyset} 2^{-l/2} |\langle f, \varphi_s\rangle| \beta_{j_0}(J^D(t,s)) \left[\int_{\mathbb{R}} g(x_1,\cdot)\delta(h_t(x_1,\cdot))dx_1\right]_{2J^D(t,s)} \\
&= \sum_{s:s\cap\Gamma_t\neq\emptyset} 2^{-l/2} w_s^{-1/2} |\langle f, \varphi_s\rangle| \beta_{j_0}(J^D(x,s)) w_s^{1/2} \left[\int_{\mathbb{R}} g(x_1,\cdot)\delta(h_t(x_1,\cdot))dx_1\right]_{2J^D(t,s)} \\
&\lesssim j_0^{3/2} \|F_t(\cdot)\|_p \left\| \int_{\mathbb{R}} g(x_1,\cdot)\delta(h_t(x_1,\cdot))dx_1 \right\|_{p'} \\
&\lesssim j_0^{3/2} \|F_t(\cdot)\|_p \left(\int_{\mathbb{R}_2} |g(x)|^{p'} \delta(h_t(x))dx\right)^{1/p'}
\end{aligned} \tag{3.3.44}$$

Hence what remains is to prove the following

Claim 3.20 *Under the above notations, we have*

$$\|F_t(\cdot)\|_p \lesssim \left(\int_{\mathbb{R}^2} \left(\sum_{k\in\mathbb{Z}} \sum_{s\in\mathcal{U}_{k,\omega}} |\langle f, \varphi_s\rangle|^2 \chi_s^2(x)\right)^{p/2} \delta(h_t(x))dx\right)^{1/p}. \tag{3.3.45}$$

Proof of Claim 3.20: By the square function estimate, we obtain

$$\|F_t\|_p \lesssim \left\| \left(\sum_{s:s\cap\Gamma_t\neq\emptyset} 2^{-l} w_s^{-2} \langle f, \varphi_s\rangle^2 \mathbb{1}_{J^D(t,s)}(\cdot)\right)^{1/2} \right\|_p. \tag{3.3.46}$$

For the right hand side of (3.3.45), again we use the new coordinate system $(v_t, v_t^\perp)$ and denote $x = x_1 v_t + x_2 v_t^\perp$ as $x = (x_1, x_2)$ for the sake of simplicity. Then the right hand side of (3.3.45) becomes

$$
\begin{aligned}
&\left(\int_{\mathbb{R}^2}\left(\sum_{k\in\mathbb{Z}}\sum_{s\in\mathcal{U}_{k,\omega}}|\langle f,\varphi_s\rangle|^2\chi_s^2(x_1,x_2)\right)^{p/2}\delta(h_t(x_1,x_2))dx_1dx_2\right)^{1/p}\\
&=\left(\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\sum_{k\in\mathbb{Z}}\sum_{s\in\mathcal{U}_{k,\omega}}|\langle f,\varphi_s\rangle|^2\chi_s^2(x_1,x_2)\delta(h_t(x_1,x_2))dx_1\right)^{p/2}dx_2\right)^{1/p}.
\end{aligned}
\tag{3.3.47}
$$

If we compare the right hand side of (3.3.46) and (3.3.47), we observe that the following pointwise estimate in x_2 will finish the proof of the claim: for any $x_2\in\mathbb{R}$ and any tile s such that $s\cap\Gamma_t\neq\emptyset$, we have

$$2^{-l}w_s^{-2}\langle f,\varphi_s\rangle^2\mathbb{1}_{J^D(t,s)}(x_2)\lesssim\int_{\mathbb{R}}|\langle f,\varphi_s\rangle|^2\chi_s^2(x_1,x_2)\delta(h_t(x_1,x_2))dx_1. \tag{3.3.48}$$

But this follows easily from the definition of the function χ_s. Thus we have finished the proof of Claim 3.20. □

3.3.3 Proof of Lemma 3.16

As we are fixing t and trying to prove pointwise estimate for $x\in\Gamma_t$, we could always pretend that the vector field is constantly equal to v_t on the whole plane. That is to say, if we define

$$\phi_s^t(x):=\int_{\mathbb{R}}\varphi_s(x-tv_t)\check{\psi}_{k-l}(t)dt,\forall x\in\mathbb{R}^2, \tag{3.3.49}$$

we will have

$$\phi_s^t(x)=\phi_s(x),\forall x\in\Gamma_t, \tag{3.3.50}$$

and the advantage is that the vector field becomes the constant vector field v_t. In the following, we will stick to ϕ_s^t instead of ϕ_s.

For a tile s of dimension $w_s\times l_s$ with

$$l_s=2^l\cdot w_s, \tag{3.3.51}$$

for a point $x\in\Gamma_t\cap s$ with

$$v_t^{\perp}\in\omega_{s,2}, \tag{3.3.52}$$

we want to show that

$$|\phi_s^t(x) - \tilde{P}_k \phi_s^t(x)| \lesssim \sum_{j_0 \in \mathbb{N}} \frac{2^{-3l/2} \cdot w_s^{-1} \beta_{j_0}(J^D(t,s))}{< j_0 >^N}. \tag{3.3.53}$$

To proceed, we again turn to the new coordinate system $(v_t, v_t^{\perp})$, and write

$$x \to x_1 v_t + x_2 v_t^{\perp}. \tag{3.3.54}$$

By the definition of the operator $\tilde{P}_k$, the left hand side of (3.3.53) is equal to

$$\phi_s^t(x_1, x_2) - \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \phi_s^t(y_1, y_2) \delta(h_t(y_1, y_2)) dy_1 \right] \psi_k(x_2 - y_2) dy_2. \tag{3.3.55}$$

We approximate $\Gamma_t \cap s$ by the line of the "average slope" in the definition of Jones' β-number, and call it $l_{s,t}$. Moreover, we define another auxiliary function $L_{s,t}$ associated to the line $l_{s,t}$ in a similar way to h_t:

$$\text{If for some } y \in \Gamma_t \text{ we have } z - y = d \cdot v_t, \text{ then we set } L_{s,t}(z) = d. \tag{3.3.56}$$

The crucial observation is that

$$\begin{aligned} &\int_{\mathbb{R}} \left[\int_{\mathbb{R}} \phi_s^t(y_1, y_2) \delta(L_{s,t}(y_1, y_2)) dy_1 \right] \psi_k(x_2 - y_2) dy_2 \\ &= \int_{\mathbb{R}} \phi_s^t(y_1, x_2) \delta(L_{s,t}(y_1, x_2)) dy_1. \end{aligned} \tag{3.3.57}$$

Substitute the above identity into (3.3.55) to obtain

$$\begin{aligned} &\phi_s^t(x_1, x_2) - \int_{\mathbb{R}} \phi_s^t(y_1, x_2) \delta(L_{s,t}(y_1, x_2)) dy_1 ... \\ &... - \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \phi_s^t(y_1, y_2) \left(\delta(h_t(y_1, y_2)) - \delta(L_{s,t}(y_1, y_2)) \right) dy_1 \right] \psi_k(x_2 - y_2) dy_2. \end{aligned} \tag{3.3.58}$$

Notice that for $x = (x_1, x_2) \in \Gamma_t$, we have

$$\phi_s^t(x_1, x_2) = \int_{\mathbb{R}} \phi_s^t(y_1, x_2) \delta(h_t(y_1, x_2)) dy_1, \tag{3.3.59}$$

by substituting which into (3.3.58) we obtain

$$\begin{aligned}&\int_{\mathbb{R}} \phi_s^t(y_1, x_2) \left[\delta(h_t(y_1, x_2)) - \delta(L_{s,t}(y_1, x_2))\right] dy_1 ...\\&... - \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \phi_s^t(y_1, y_2) \left(\delta(h_t(y_1, y_2)) - \delta(L_{s,t}(y_1, y_2))\right) dy_1\right] \psi_k(x_2 - y_2) dy_2.\end{aligned} \tag{3.3.60}$$

Observe that the latter term in the above expression is just a Littlewood-Paley projection of the former term, hence it should be expected that these two terms can be handled in a similar way. In Subsection 2.4.3 it has indeed been shown to be this case, hence in the following we will focus on the former term of (3.3.60).

By the definition of ϕ_s^t in (3.3.49), we obtain

$$\begin{aligned}&\int_{\mathbb{R}} \phi_s^t(y_1, x_2) \left[\delta(h_t(y_1, x_2)) - \delta(L_{s,t}(y_1, x_2))\right] dy_1\\&= \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi_s(y_1 - t, x_2) \check{\psi}_{k-l}(t) dt \left[\delta(h_t(y_1, x_2)) - \delta(L_{s,t}(y_1, x_2))\right] dy_1.\end{aligned} \tag{3.3.61}$$

If we denote

$$d := h_t(y_1, x_2) - L_{s,t}(y_1, x_2), \tag{3.3.62}$$

then the right hand side of (3.3.61) turns to

$$\begin{aligned}&\int_{\mathbb{R}} \int_{\mathbb{R}} \left(\varphi_s(y_1 - t, x_2) - \varphi_s(y_1 + d - t, x_2)\right) \check{\psi}_{k-l}(t) dt \delta(h_t(y_1, x_2)) dy_1\\&= \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi_s(y_1 - t, x_2) \left(\check{\psi}_{k-l}(t) - \check{\psi}_{k-l}(t + d)\right) dt \delta(h_t(y_1, x_2)) dy_1.\end{aligned} \tag{3.3.63}$$

Hence by the definition of Jones' beta numbers that

$$|d| \lesssim w_s \cdot \beta_0(J^D(t, s)), \tag{3.3.64}$$

and by applying the fundamental theorem of calculus to $\check{\psi}_{k-l}$, we conclude the desired estimate in Lemma 3.16. □

Chapter 4

A geometric proof of Bourgain's L^2 estimate of the maximal operator along analytic vector fields

In this chapter we present a geometric proof of Bourgain's L^2 bounds of the maximal operator along analytic vector fields, which is Theorem 1.6. Recall that for a unit vector field $v : \mathbb{R}^2 \to S^1$, for a small constant $\epsilon_0 > 0$, the maximal operator associated to v and truncated at ϵ_0 is defined by

$$M_{v,\epsilon_0} f(x) := \sup_{0<\epsilon\leq\epsilon_0} \frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} |f(x + tv(x))| dt. \tag{4.0.1}$$

It is a result due to Bourgain [7] that for every analytic vector field v, there exists ϵ_0 depending on v such that M_{v,ϵ_0} is bounded on $L^2(\mathbb{R}^2)$. As we are doing a truncation in the definition of the maximal operator (4.0.1), it suffices to prove that for any bounded open set $\Omega \subset \mathbb{R}^2$, M_{v,ϵ_0} is bounded on $L^2(\Omega)$.

To prove Theorem 1.6, Bourgain reduced the analyticity assumption on the vector field to the following geometric one: for $x \in \Omega$ and t small enough, define the function

$$\omega_x(t) = |\det[v(x + tv(x)), v(x)]| \,. \tag{4.0.2}$$

We assume that

$$\left| \{t \in [-\epsilon, \epsilon] : \omega_x(t) < \tau \sup_{-\epsilon\leq s\leq\epsilon} \omega_x(s)\} \right| \leq C_0 \tau^{c_0} \epsilon, \tag{4.0.3}$$

for all $0 < \tau < 1, 0 < \epsilon \leq \epsilon_0$, where $0 < c_0, C_0 < \infty$ are constants independent of the point $x \in \Omega$.

It is shown in [7] that Theorem 1.6 can be reduced to the following

Theorem 4.1 ([7]) *If v is C^1 and satisfies the condition* (4.0.3)*, then M_{v,ϵ_0} is bounded on $L^2(\Omega)$.*

Bourgain's proof for Theorem 4.1 is not entirely geometric, particularly the key Lemma 3.28, where he used the polar coordinates and applied Schur's Lemma to get the desired L^2 bounds.

The goal of this chapter is to give a geometric proof of Theorem 4.1. The idea is to use the time-frequency decomposition initiated by Lacey and Li in the setting of the Hilbert transform along vector fields in [20] and [21], and further developed by Bateman in [4], Bateman and Thiele in [5]. However, the proof below is free of the time-frequency analysis techniques.

4.1 Reduction to a smooth cut-off

By a renormalisation, we assume further that $\|v\|_{C^1} \leq 1$, and $\Omega = B_{\epsilon_0}(0)$, which is the ball of radius $\epsilon_0 \ll 1$ centered at origin. Moreover, as we are only concerned with the truncated maximal operator, we can w.l.o.g. assume that the vector field is periodic in both horizontal and vertical directions with each periodicity being $3 \cdot \epsilon_0$, and that the vector field always points in the two-ended cone which forms an angle less than $\pi/10$ with the horizontal axis. In the following, we will denote this cone by Γ_0.

Choose $\alpha : \mathbb{R} \to \mathbb{R}$ to be a proper smooth bump function such that the support of $\hat{\alpha}$ lies on $[-1, 1]$. For $0 < \epsilon \leq \epsilon_0$, define

$$A_\epsilon f(x) := \int_{\mathbb{R}} f(x + \epsilon t v(x))\alpha(t)dt. \tag{4.1.1}$$

It is not difficult to see that the operator M_{v,ϵ_0} is essentially equivalent to

$$\sup_{j\in\mathbb{N}, 2^{-j}\leq\epsilon_0} |A_{2^{-j}} f(x)|, \tag{4.1.2}$$

which will still be called M_{v,ϵ_0}. Moreover we will write A^j to stand for $A_{2^{-j}}$ for the sake of simplicity. Hence in the rest of the paper, we will focus on the following operator

$$M_{v,\epsilon_0} f(x) := \sup_{j\in\mathbb{N}, 2^{-j}\leq\epsilon_0} |A^j f(x)|. \tag{4.1.3}$$

In the end, we just need to observe that to prove (4.1.3), it suffices to prove the following spatially localised version:

$$\|M_{v,\epsilon_0} f\|_{L^2(B_{\epsilon_0}(0))} \lesssim \|f\|_{L^2}, \tag{4.1.4}$$

due to the fact that we are truncating the maximal operator at the scale ϵ_0.

4.2 Bourgain's high-low frequency decomposition

We linearise the maximal operator in (4.1.3): take a measurable function $J : \mathbb{R}^2 \to \mathbb{N}$ such that

$$M_{v,\epsilon_0} f(x) \sim \left| A^{J(x)} f(x) \right|. \tag{4.2.1}$$

For a point $x \in \Omega = B_{\epsilon_0}(0)$, let $R_{x,j}$ be the rectangle with center x, orientation $v(x)$, length 2^{-j} in direction $v(x)$ and width

$$\delta(R_{x,\epsilon}) = 2^{-j} \cdot \sup_{|t|<2^{-j}} \omega_x(t). \tag{4.2.2}$$

Especially we denote

$$\delta(x) := 2^{-J(x)} \cdot \sup_{|t|<2^{-J(x)}} \omega_x(t). \tag{4.2.3}$$

Choose a measurable function $K : \mathbb{R}^2 \to \mathbb{N}$ such that

$$\delta(x) \sim 2^{-K(x)}, \forall x \in \mathbb{R}^2. \tag{4.2.4}$$

Do an isotropic Littlewood-Paley decomposition for the function f, and write

$$f = \sum_{k \in \mathbb{Z}} P_k f. \tag{4.2.5}$$

This turns the operator into

$$\sum_{k \in \mathbb{Z}} A^{J(x)} P_k f(x). \tag{4.2.6}$$

Bourgain's idea is to split the function into two parts, the high frequency part and the low frequency part, in the following way:

$$\sum_{k \in \mathbb{Z}} A^{J(x)} P_k f(x) = \sum_{k \in \mathbb{Z}, k \geq K(x)} A^{J(x)} P_k f(x) + \sum_{k \in \mathbb{Z}, k < K(x)} A^{J(x)} P_k f(x). \tag{4.2.7}$$

For the latter part, i.e. the low frequency part, Bourgain's proof is already geometric, see Lemma 4.12 and Lemma 5.7 in [7]. Hence the main task for us is to bound the former part, i.e. the high frequency part, by a geometric argument.

Remark 4.2 *The estimate of the above high frequency part is done in Lemma 3.28 in [7] by analytic* *methods.*

We proceed with the estimate of the high frequency part: First we write

$$\sum_{k\in\mathbb{Z},k\geq K(x)} A^{J(x)} P_k f(x) = \sum_{l\in\mathbb{N}_0} A^{J(x)} P_{K(x)+l} f(x). \tag{4.2.8}$$

Then by the triangle inequality, it suffices to prove that

$$\|A^{J(x)} P_{K(x)+l} f(x)\|_2 \lesssim 2^{-\mu l} \|f\|_2, \tag{4.2.9}$$

for some $\mu > 0$, with constant being independent of $l \in \mathbb{N}$.

Notice that the above estimate is still of maximal type, and we want to get rid of the linearisation by replacing the l^∞ norm by an l^2 norm. To do this, we need to introduce several notations. For $j, k \in \mathbb{N}$, define

$$\Omega_{j,k} := \{x \in \Omega | 2^{-j} \cdot \sup_{|t|<2^{-j}} \omega_x(t) \sim 2^{-k}\}. \tag{4.2.10}$$

For a real analytic vector field, either the integral curves are straight lines, or for each $j \in \mathbb{N}$, the complement of the set $\cup_k \Omega_{j,k}$ has measure zero. Hence it is no restriction to assume for each $j \in \mathbb{N}$ that

$$\Omega = \bigcup_{k\in\mathbb{N}} \Omega_{j,k}. \tag{4.2.11}$$

It is also clear that for a fixed $k \in \mathbb{N}$, the $\Omega_{j,k}$ for different j are essentially disjoint.

Hence for a fixed x,

$$|A^{J(x)} P_{K(x)+l} f(x)| \lesssim \sup_{j\in\mathbb{N}} \left(\sum_{k\in\mathbb{N}} \left| A^j P_{k+l} f \right|^2 \mathbb{1}_{\Omega_{j,k}} \right)^{1/2}. \tag{4.2.12}$$

We replace the sup norm by the l^2 norm to obtain

$$|A^{J(x)} P_{K(x)+l} f(x)| \lesssim \left(\sum_{j\in\mathbb{N}} \sum_{k\in\mathbb{N}} \left| A^j P_{k+l} f \right|^2 \mathbb{1}_{\Omega_{j,k}} \right)^{1/2}. \tag{4.2.13}$$

Taking the L^2 norm of (4.2.13), we obtain

$$\sum_{j\in\mathbb{N}}\sum_{k\in\mathbb{N}}\int \left|A^j P_{k+l} f\right|^2 \mathbb{1}_{\Omega_{j,k}}. \tag{4.2.14}$$

Recall that $R_{x,j}$ denotes the rectangle with center x, length 2^{-j} and width

$$\delta(R_{x,j}) := 2^{-j} \sup_{|t|<2^{-j}} \omega_x(t). \tag{4.2.15}$$

In the following we will cover $\Omega_{j,k}$ with rectangles $\{R_m = R_{x_m,j}\}_{m\in\mathbb{N}}$ satisfying the following two conditions

$$\begin{aligned} &(i)\ \ \delta(R_m) = 2^{-k}; \\ &(ii)\ \ \text{the center of } R_m \text{ does not belong to } R_1 \cup \ldots \cup R_{m-1}. \end{aligned} \tag{4.2.16}$$

Hence for fixed j and k,

$$\int \left|A^j P_{k+l} f\right|^2 \mathbb{1}_{\Omega_{j,k}} \lesssim \sum_{m\in\mathbb{N}} \int \left|A^j P_{k+l} f\right|^2 \mathbb{1}_{R_m}. \tag{4.2.17}$$

Indeed, the above covering of $\Omega_{j,k}$ is a "valid" covering, i.e. a covering without much overlapping. To be precise, if we define

$$\Omega'_{j,k} := \bigcup_{x\in\Omega_{j,k}} 2\cdot R_{x,j}, \tag{4.2.18}$$

then it has been proved by Bourgain in [7] (see the following Lemma 4.7) that

$$\|\sum_j \mathbb{1}_{\Omega'_{j,k}}\|_\infty \lesssim 1. \tag{4.2.19}$$

In the following, when estimating the right hand side of (4.2.17), we will need several other geometric properties like (4.2.19). Hence we organise all them together in the next section.

4.3 Geometric properties of the rectangles

Lemma 4.3 (Lemma 4.1 in [7]) *Let x' be in the rectangle $R_{x,j}$, then*

$$\delta(R_{x,j}) \sim \delta(R_{x',j}), \tag{4.3.1}$$

and $R_{x,j}$ is contained in a multiple of $R_{x',j}$ and vice versa.

Lemma 4.4 (Lemma 4.6 in [7]) *Assume*

$$2 \cdot R_{x,j} \cap 2 \cdot R_{x',j'} \neq \emptyset, \tag{4.3.2}$$

and $2^{-j'} \lesssim 2^{-j}$. Then

$$R_{x',j'} \subset 4 \cdot R_{x,j} \tag{4.3.3}$$

and

$$\frac{\delta(R_{x,j})}{2^{-j}} \gtrsim \frac{\delta(R_{x',j'})}{2^{-j'}}, \tag{4.3.4}$$

i.e. larger rectangle has larger eccentricity.

Lemma 4.5 *Assume that*

$$R_{x,j} \cap R_{x',j+j_0} \neq \emptyset, \tag{4.3.5}$$

for some $j_0 \in \mathbb{N}_0$. Then there exists a constant $a_0 > 1$ such that

$$\frac{\delta(R_{x,j})}{2^{-j}} \lesssim (2^{j_0})^{a_0} \cdot \frac{\delta(R_{x',j+j_0})}{2^{-j-j_0}}. \tag{4.3.6}$$

Remark 4.6 *Compared with Lemma 4.4, this lemma says that the growth of the eccentricity of the rectangle with respect to the length can only be polynomial.*

Proof of Lemma 4.5: this follows easily from Lemma 4.3 and the following doubling estimate (3.20) in [7]

$$\frac{\delta(R_{x,j})}{2^{-j}} \leq C \cdot \frac{\delta(R_{x',j+1})}{2^{-j-1}}, \tag{4.3.7}$$

for some constant $C > 0$. □

Lemma 4.7 (Lemma 4.7 in [7]) *Let $\{R_{x_i,j_i}\}_{i\in\mathbb{N}_0}$ be a sequence of rectangles and $\delta > 0$ such that*

$$\begin{aligned} &(i)\ \delta(R_{x_i,j_i}) \sim \delta; \\ &(ii)\ x_{i+1} \text{ does not belong to } R_{x_0,j_0} \cup \ldots \cup R_{x_i,j_i}, \forall i. \end{aligned} \tag{4.3.8}$$

Then

$$\| \sum_{i\in\mathbb{N}_0} \mathbb{1}_{2\cdot R_{x_i,j_i}} \|_\infty \lesssim 1. \tag{4.3.9}$$

We will also need the following generalised version of the above lemma.

Lemma 4.8 *Under the same assumptions as in Lemma 4.7, there exists a constant* $b_0 > 0$ *such that for any* $N \in \mathbb{N}_0$*, we have*

$$\| \sum_{i\in\mathbb{N}_0} \mathbb{1}_{R^{p,q}_{x_i,j_i}} \|_\infty \lesssim (p+q+1)^{b_0}, \forall p, q \in \mathbb{N}, \tag{4.3.10}$$

where $R^{p,q}_{x_i,j_i}$ *is obtained by dilating the length of* R_{x_i,j_i} *to* p *times, and the width to* q *times.*

Remark 4.9 *The above lemma says that when we enlarge the rectangles* R_{x_i,j_i}*, the overlapping can only be polynomially growing.*

Proof of Lemma 4.8: Denote

$$\tilde{q} := \max\{q, p^{a_0+1}\}, \tag{4.3.11}$$

where a_0 is the constant in (4.3.6).For the rectangle $R^{p,q}_{x_i,j_i}$, we further enlarge it to be of width $\tilde{q}\cdot\delta$. Next, we will dilate the length to $\tilde{p}_i \cdot 2^{-j_i}$ such that

$$\sup_{|t|\le\tilde{p}_i\cdot 2^{-j_i}} \omega_{x_i}(t) \sim \frac{\tilde{q}}{\tilde{p}_i}\cdot\frac{\delta}{2^{-j_i}}. \tag{4.3.12}$$

By Lemma 4.4, it is not difficult to see that

$$\tilde{p}_i \le \tilde{q} \lesssim \tilde{p}_i^{a_0+1}, \tag{4.3.13}$$

uniformly in i.

Our goal now is to show that

$$\| \sum_i \mathbb{1}_{R^{\tilde{p}_i,\tilde{q}}_{x_i,j_i}} \|_\infty \lesssim \tilde{q}^{b_0}, \tag{4.3.14}$$

for some b_0 to be determined later. Suppose that the L^∞ norm on the left hand side of the above expression is attained at the point O. Moreover, let $\mathcal{R}_O$ denote the collection of rectangles containing the point O, and w.l.o.g. we assume that

$$\mathcal{R}_O = \{R^{\tilde{p}_i,\tilde{q}}_{x_i,j_i}\}_{0\le i\le N}, \tag{4.3.15}$$

for some $N \in \mathbb{N}_0$. Then (4.3.14) is equivalent to proving

$$N \lesssim \tilde{q}^{b_0}. \tag{4.3.16}$$

By the definition of the rectangles $R^{\tilde{p}_i,\tilde{q}}_{x_i,j_i}$ and Lemma 4.3, it is not difficult to see that all the rectangles in $\mathcal{R}_O$ have comparable lengths. Indeed, up to a constant dilation factor, any of these rectangles is contained in another. Hence

$$\bigcup_{0\le i\le N} R^{\tilde{p}_i,\tilde{q}}_{x_i,j_i} \subset 4R^{\tilde{p}_0,\tilde{q}}_{x_0,j_0}. \tag{4.3.17}$$

Moreover, by the upper bound on $\tilde{p}_i$ in (4.3.13), we can also obtain that

$$l(R_{x_i,j_i}) \gtrsim l(R^{\tilde{p}_0,\tilde{q}}_{x_0,j_0})/\tilde{q}, \tag{4.3.18}$$

where for some rectangle R, $l(R)$ is used to denote its length. Hence by the assumption that the center x_i of R_{x_i,j_i} is not contained in

$$R_{x_0,j_0} \cup \ldots \cup R_{x_{i-1},j_{i-1}} \tag{4.3.19}$$

for all i, we obtain easily the estimate (4.3.16) for some constant b_0 depending only on a_0. So far we have finished the proof of Lemma 4.8. $\square$

4.4 Estimate on each rectangle

In this section, we will give an estimate of each single term from the summation on the right hand side of (4.2.17), i.e.

$$\int |A^j P_{k+l} f|^2 \mathbb{1}_{R_m}, \tag{4.4.1}$$

for a fixed $R_m = R_{x_m,j}$. Recall that we have assumed that the vector field points in the cone Γ_0, which is the two-ended cone forming an angle less than $\pi/10$ with the horizontal axis. If we denote by P_{Γ_0} the frequency projection operator for the cone Γ_0 (as in (2.2.8)), it is not difficult to see that

$$A^j P_{k+l} P_{\Gamma_0} f \equiv 0. \tag{4.4.2}$$

Hence in the following we will only be concerned with the frequency in the cone Γ_0^c. Moreover, for the sake of simplicity, we will always identify P_{k+l} with $P_{k+l}P_{\Gamma_0^c}$.

Now we use the time-frequency decomposition from Section 2.4 to write the function $A^j P_{k+l}$ into a model sum. For fixed j, k and l, we will denote

$$\theta := k + l - j. \tag{4.4.3}$$

By using the notations from (2.4.1) to (2.4.13), the frequency localised function $A^j P_{k+l} f$ can be passed to the model sum

$$\sum_{\omega \in \mathcal{D}_\theta} \sum_{s \in \mathcal{U}_{k+l,\omega}} \langle f, \varphi_s \rangle A^j \varphi_s.$$

Hence to bound the expression (4.4.1), it suffices to bound

$$\int \left| \sum_{\omega \in \mathcal{D}_\theta} \sum_{s \in \mathcal{U}_{k+l,\omega}} \langle f, \varphi_s \rangle A^j \varphi_s \right|^2 \mathbb{1}_{R_m}. \tag{4.4.4}$$

By Lemma 2.14, we have that the last expression is equal to

$$\sum_{\omega \in \mathcal{D}_\theta} \int \left| \sum_{s \in \mathcal{U}_{k+l,\omega}} \langle f, \varphi_s \rangle A^j \varphi_s \right|^2 \mathbb{1}_{R_m}. \tag{4.4.5}$$

We will focus on (4.4.5) in the following two subsections.

4.4.1 Estimate on each rectangle by ignoring the tails of the wavelet functions

In this part, we will only show the idea of how to bound the term (4.4.5), or in another word, we will ignore the tails of the wavelet functions and the function α in the definition of A^j in (4.1.1), and always assume that they have compact support in both space and frequency.

Under the above simplification, the expression in (4.4.5) turns to

$$\sum_{\omega \in \mathcal{D}_\theta} \sum_{s \in \mathcal{U}_{k+l,\omega}} |\langle f, \varphi_s \rangle|^2 \int \left| A^j \varphi_s \right|^2 \mathbb{1}_{R_m}. \tag{4.4.6}$$

Take a point $x \in R_m$, for a tile $s \in \mathcal{U}_{k+l,\omega}$ for some $\omega \in \mathcal{D}_\theta$, we observe that in order for $A^j \varphi_s(x)$ not to vanish, we must have $\omega \subset 3 \cdot EX(R_m)$ as by Lemma 4.3 we know that $v(x) \in 2 \cdot EX(R_m)$ for any $x \in R_m$. Here $EX(R_m)$ is the uncertainty interval of the rectangle R_m given in Definition

2.8. This, together with the fact that both R_m and s have length 2^{-j}, implies that

$$s \subset 4 \cdot R_m, \tag{4.4.7}$$

for those tiles s such that $A^j \varphi_s$ is not identically zero.

Claim 4.10 *There exists $\mu > 0$ such that*

$$\int \left|A^j \varphi_s\right|^2 \mathbb{1}_{R_m} \lesssim 2^{-\mu l}, \tag{4.4.8}$$

with the constant being independent of s.

By the above claim, the expression in (4.4.6) can be further bounded by

$$\sum_{\omega \in \mathcal{D}_\theta} \sum_{s \in \mathcal{U}_{k+l,\omega}, s \subset 4 \cdot R_m} 2^{-\mu l} \cdot |\langle f, \varphi_s \rangle|^2 \lesssim 2^{-\mu l} \|\mathbb{1}_{R_m} \cdot P_{k+l} f\|_2^2. \tag{4.4.9}$$

The next step is to sum over m, j and k:

$$\begin{aligned} \sum_{j,k} \sum_{m} \|\mathbb{1}_{R_m} \cdot P_{k+l} f\|_2^2 &\lesssim 2^{-\mu l} \sum_{j,k} \|\mathbb{1}_{\Omega'_{j,k}} \cdot P_{k+l} f\|_2^2 \\ &\lesssim 2^{-\mu l} \sum_{k} \|P_{k+l} f\|_2^2 \lesssim 2^{-\mu l} \|f\|_2^2, \end{aligned} \tag{4.4.10}$$

where we have used the disjointness property (4.2.19). Hence for the model problem, what remains is

"Proof" of Claim 4.10: We can w.l.o.g. assume that there exists a point $x_0 \in s$ such that

$$v(x_0) \in \omega_s, \tag{4.4.11}$$

as otherwise $A^j \varphi_s$ will be identically zero. By a proper translation and rotation, we can assume that $x_0 = (0, 0)$ and $v(x_0) = (1, 0)$.

Now we look at the direction of the vector field for the points on the line segment

$$\{(x_1, x_2) : x_2 = 0\} \cap s. \tag{4.4.12}$$

By the assumption on the rectangle R_m we know that

$$\sup_{|t|\le 2^{-j}} w_{x_0}(t) = \sup_{|t|\le 2^{-j}} |\det[v(x_0+tv(x_0)), v(x_0)]| \sim 2^{-k+j}. \tag{4.4.13}$$

Notice that $|\omega_s| = 2^{-k-l+j}$, hence in order for $A^j\varphi_s$ not to vanish at a point $x \in s \cap \{(x_1, x_2) : x_2 = 0\}$, we must have

$$|\det[v(x), v(x_0)]| = w_{x_0}(x \cdot v(x_0)) \lesssim 2^{-k-l+j}. \tag{4.4.14}$$

By taking $\tau = 2^{-l}$ in the condition (4.0.3) we obtain

$$\left| t \in [-2^{-j}, 2^{-j}] : w_{x_0}(t) < 2^{-l} \sup_{|t|\le 2^{-j}} w_{x_0}(t) \right| \le C_0 2^{-c_0 l} \cdot 2^{-j}, \tag{4.4.15}$$

which further implies that

$$\left|\{(x_1, 0) \in s : A^j\varphi_s(x_1, 0) \neq 0)\}\right| \lesssim 2^{-c_0 l} \cdot 2^{-j}. \tag{4.4.16}$$

So far we have proved that on one line segment, the non-vanishing output has relatively small measure. In the next, we want to show that this indeed holds true for all the points in the tile s, namely

$$|\{x \in s : A^j\varphi_s(x) \neq 0\}| \lesssim 2^{-c_0 l}|s|. \tag{4.4.17}$$

This, combined with the trivial estimate

$$\|A^j\varphi_s\|_\infty \lesssim |s|^{-1/2}, \tag{4.4.18}$$

concludes the proof of Claim 4.10.

We turn to the proof of (4.4.17): for $|x_2| \le 2^{-k-l+2}$, consider the line segment

$$L_{x_2} := \{(0, x_2) + t \cdot v(0, x_2) : |t| \le 2^{-j+2}\}. \tag{4.4.19}$$

First by the C^1 assumption on the vector field, we know that

$$v(0, x_2) \in 2 \cdot \omega_s, \forall |x_2| \le 2^{-k-l+2}. \tag{4.4.20}$$

Then by the same argument as before, we obtain that

$$|\{x \in L_{x_2} : A^j \varphi_s(x) \neq 0\}| \lesssim 2^{-c_0 l} \cdot 2^{-j}, \tag{4.4.21}$$

for each $|x_2| \leq 2^{-k-l+2}$. Hence by Fubini's theorem (which can be applied due to the C^1 assumption on the vector field), we obtain

$$\begin{aligned} &|\{x \in s : A^j \varphi_s(x) \neq 0\}| \\ &= \int_{-2^{-k-l+2}}^{2^{-k-l+2}} |\{x \in L_{x_2} : A^j \varphi_s(x) \neq 0\}| dx_2 \lesssim 2^{-c_0 l} \cdot 2^{-k-l-j}. \end{aligned} \tag{4.4.22}$$

Hence we have finished the proof of (4.4.17).

4.4.2 The full estimate on each rectangle

In this part, we will make the above heuristic argument rigorous, i.e. we will also take care of the tails of the wavelet functions. For fixed j, k, l and m, we want to bound the following

$$\int \left|A^j P_{k+l} f\right|^2 \mathbb{1}_{R_m} = \sum_{\omega \in \mathcal{D}_\theta} \int \left| \sum_{s \in \mathcal{U}_{k+l,\omega}} \langle f, \varphi_s \rangle A^j \varphi_s \right|^2 \mathbb{1}_{R_m}. \tag{4.4.23}$$

For $p, q \in \mathbb{Z}$, we denote by $\vec{R}_m^{p,q}$ the translation of the rectangle R_m by (p, q) units, i.e.

$$\vec{R}_m^{p,q} = R_m + p \cdot 2^{-j} v(x_m) + q \cdot 2^{-k} v^\perp(x_m), \tag{4.4.24}$$

where x_m denotes the center of R_m and $v(x_m)$ is the value of the vector field at the point x_m which is parallel to the long side of R_m.

Hence for one fixed $\omega \in \theta$, we have

$$\int \left| \sum_{s \in \mathcal{U}_{k+l,\omega}} \langle f, \varphi_s \rangle A^j \varphi_s \right|^2 \mathbb{1}_{R_m} = \int \left| \sum_{p,q \in \mathbb{Z}} \sum_{s \in \mathcal{U}_{k+l,\omega}, s \subset \vec{R}_m^{p,q}} \langle f, \varphi_s \rangle A^j \varphi_s \right|^2 \mathbb{1}_{R_m}. \tag{4.4.25}$$

By Minkowski's inequality, the right hand side of the above display can be bounded by

$$\left(\sum_{p,q \in \mathbb{Z}} \left(\int \left| \sum_{s \in \mathcal{U}_{k+l,\omega}, s \subset \vec{R}_m^{p,q}} \langle f, \varphi_s \rangle A^j \varphi_s \right|^2 \mathbb{1}_{R_m} \right)^{1/2} \right)^2. \tag{4.4.26}$$

Lemma 4.11 *For any large $M \in \mathbb{N}_0$, there exists a constant C_M such that*

$$\int \left| \sum_{s \in \mathcal{U}_{k+l,\omega}, s \subset \vec{R}_m^{p,q}} \langle f, \varphi_s \rangle A^j \varphi_s \right|^2 \mathbb{1}_{R_m} \lesssim \frac{2^{-\mu l}}{(|p|+|q|+1)^M} \sum_{s \in \mathcal{U}_{k+l,\omega}, s \subset \vec{R}_m^{p,q}} |\langle f, \varphi_s \rangle|^2, \tag{4.4.27}$$

where μ is the same as the one in Claim 4.10.

We substitute the estimate in Lemma 4.11 into (4.4.26) to obtain

$$\begin{aligned} &\left(\sum_{p,q \in \mathbb{Z}} \left(\frac{2^{-\mu l}}{(|p|+|q|+1)^M} \sum_{s \in \mathcal{U}_{k+l,\omega}, s \subset \vec{R}_m^{p,q}} |\langle f, \varphi_s \rangle|^2 \right)^{1/2} \right)^2 \\ &\lesssim \sum_{p,q \in \mathbb{Z}} \frac{2^{-\mu l}}{(|p|+|q|+1)^M} \sum_{s \in \mathcal{U}_{k+l,\omega}, s \subset \vec{R}_m^{p,q}} |\langle f, \varphi_s \rangle|^2, \end{aligned} \tag{4.4.28}$$

where the exact value of M might vary from line to line. Hence we have obtained

$$\int \left| A^j P_{k+l} f \right|^2 \mathbb{1}_{R_m} \lesssim \sum_{\omega \in \mathcal{D}_\theta} \sum_{p,q \in \mathbb{Z}} \frac{2^{-\mu l}}{(|p|+|q|+1)^M} \sum_{s \in \mathcal{U}_{k+l,\omega}, s \subset \vec{R}_m^{p,q}} |\langle f, \varphi_s \rangle|^2, \tag{4.4.29}$$

which is the estimate on one single rectangle that we are aiming at.

Proof of Lemma 4.11: we will only consider the case $p = q = 0$, and the decay in p and q in the other case will simply follow from the non-stationary phase method. Hence what we need to prove becomes

$$\int \left| \sum_{s \in \mathcal{U}_{k+l,\omega}, s \subset R_m} \langle f, \varphi_s \rangle A^j \varphi_s \right|^2 \mathbb{1}_{R_m} \lesssim 2^{-\mu l} \sum_{s \in \mathcal{U}_{k+l,\omega}, s \subset R_m} |\langle f, \varphi_s \rangle|^2. \tag{4.4.30}$$

Recall that each tile s has width 2^{-k-l}, however the rectangle R_m has width 2^{-k}. This suggests that we should do a further partition of R_m into smaller rectangles which will be of the same scale as s.

We enumerate the tiles $s \subset R_m$ from above to below by $s_1, s_2, ..., s_{m'}...$, where $m' \lesssim 2^l$. Notice that

$$R_m \subset \bigcup_{m'} 2 \cdot s_{m'}. \tag{4.4.31}$$

Hence

$$\int \left| \sum_{s\in\mathcal{U}_{k+l,\omega}, s\subset R_m} \langle f, \varphi_s\rangle A^j \varphi_s \right|^2 \mathbb{1}_{R_m} \lesssim \sum_{m'} \int \left| \sum_{m''} \langle f, \varphi_{s_{m''}}\rangle A^j \varphi_{s_{m''}} \right|^2 \mathbb{1}_{2\cdot s_{m'}}. \tag{4.4.32}$$

By the Cauchy-Schwartz inequality, we bound the right hand side of the above expression by

$$\sum_{m'} \int \sum_{m''} \left| \langle f, \varphi_{s_{m''}}\rangle A^j \varphi_{s_{m''}} \right|^2 (|m' - m''| + 1)^M \mathbb{1}_{2\cdot s_{m'}}, \tag{4.4.33}$$

for some large constant M. By the same argument as in the proof of Claim 4.10, we obtain

$$|\{x \in 2 \cdot s_{m'} : A^j \varphi_{s_{m''}} \neq 0\}| \lesssim 2^{-c_0 l} |s_{m'}|. \tag{4.4.34}$$

This, together with the trivial bound

$$\|A^j \varphi_{s_{m''}}\|_{L^\infty(s_{m'})} \lesssim \frac{|s_{m''}|^{1/2}}{(|m' - m''| + 1)^M}, \tag{4.4.35}$$

implies that

$$\int \left| \langle f, \varphi_{s_{m''}}\rangle A^j \varphi_{s_{m''}} \right|^2 \mathbb{1}_{2\cdot s_{m'}} \lesssim \frac{2^{-c_0 l}}{(|m' - m''| + 1)^{2M}} |\langle f, \varphi_{s_{m''}}\rangle|^2. \tag{4.4.36}$$

We substitute the above estimate into (4.4.33) to obtain

$$\sum_{m'} \sum_{m''} \frac{2^{-c_0 l}}{(|m' - m''| + 1)^M} |\langle f, \varphi_{s_{m''}}\rangle|^2 \lesssim \sum_{m''} |\langle f, \varphi_{s_{m''}}\rangle|^2. \tag{4.4.37}$$

So far we have finished the proof of Lemma 4.11, hence the estimate on each rectangle, which is (4.4.29).

4.5 Organising all the rectangles together to finish the proof

In this part, we will organise the estimates on all the rectangles together, i.e. to finish the proof of the following estimate

$$\sum_{j,k} \sum_{m} \int \left| A^j P_{k+l} f \right|^2 \mathbb{1}_{R_m} \lesssim 2^{-\mu l} \|f\|_2^2, \tag{4.5.1}$$

for some $\mu > 0$. To do this, we substitute the estimate (4.4.29) into the left hand side of the above expression to obtain

$$\sum_{k,j}\sum_{m}\sum_{\omega\in\mathcal{D}_\theta}\sum_{p,q\in\mathbb{Z}}\frac{2^{-\mu l}}{(|p|+|q|+1)^M}\sum_{s\in\mathcal{U}_{k+l,\omega},s\subset\vec{R}_m^{p,q}}|\langle f,\varphi_s\rangle|^2, \tag{4.5.2}$$

where we are still using the notation $\theta = k + l - j$. Hence it suffices to prove that for fixed $p, q \in \mathbb{Z}$, we have

$$\sum_{k,j}\sum_{m}\sum_{\omega\in\mathcal{D}_\theta}\sum_{s\in\mathcal{U}_{k+l,\omega},s\subset\vec{R}_m^{p,q}}|\langle f,\varphi_s\rangle|^2 \lesssim (|p|+|q|+1)^{b_0}\|f\|_2^2, \tag{4.5.3}$$

where b_0 is the constant in Lemma 4.8.

Proof of the estimate (4.5.3)**:** We first fix k. For the case $p = q = 0$, for two tiles s' and s'' in the following collection

$$\bigcup_{j}\bigcup_{\omega\in\mathcal{D}_\theta}\bigcup_{m}\{s : s\in\mathcal{U}_{k+l,\omega}, s\subset R_m\}, \tag{4.5.4}$$

we either have

$$\omega_{s'}\cap\omega_{s''} = \emptyset, \tag{4.5.5}$$

or

$$s'\cap s'' = \emptyset. \tag{4.5.6}$$

Hence by the (almost) orthogonality, we obtain that

$$\sum_{j}\sum_{m}\sum_{\omega\in\mathcal{D}_\theta}\sum_{s\in\mathcal{U}_{k+l,\omega},s\subset R_m}|\langle f,\varphi_s\rangle|^2 \lesssim \|P_{k+l}f\|_2^2. \tag{4.5.7}$$

By summing over k, we get the desired estimate (4.5.3) for the case $p = q = 0$.

For the general $p, q \in \mathbb{Z}$, we no longer have (4.5.6) due to the simple fact that for two disjoint rectangles (of different scales), they might intersect after being translated by (p, q) units separately. Fortunately, Lemma 4.8 says that the intersection caused by translation can only grow polynomially in p and q.

Hence by essentially the same idea as above and by losing a factor of $(|p|+|q|+1)^{b_0}$, we obtain

$$\sum_{j}\sum_{m}\sum_{\omega\in\mathcal{D}_\theta}\sum_{s\in\mathcal{U}_{k+l,\omega},s\subset\vec{R}_m^{p,q}}|\langle f,\varphi_s\rangle|^2 \lesssim (|p|+|q|+1)^{b_0}\|P_{k+l}f\|_2^2. \tag{4.5.8}$$

Summing over k, we get the estimate (4.5.3). So far we have finished the proof of (4.5.1), hence the geometric proof of Theorem 4.1.

Bibliography

[1] P. Auscher, S. Hofmann, C. Muscalu, T. Tao and C. Thiele: Carleson measures, trees, extrapolation, and $T(b)$ theorems. Publ. Mat. 46 (2002), no. 2, 257-325.

[2] J. Azzam and R. Schul: Hard Sard: quantitative implicit function and extension theorems for Lipschitz maps. Geom. Funct. Anal. 22 (2012), no. 5, 1062-1123.

[3] M. Bateman: L^p estimates for maximal averages along one-variable vector fields in $\mathbb{R}^2$. Proc. Amer. Math. Soc. 137 (2009), no. 3, 955-963.

[4] M. Bateman: Single annulus L^p estimates for Hilbert transforms along vector fields. Rev. Mat. Iberoam. 29 (2013), no. 3, 1021-1069.

[5] M. Bateman and C. Thiele: L^p estimate for the Hilbert transform along a one variable vector field. Anal. PDE 6 (2013), no. 7, 1577-1600.

[6] F. Bernicot: Fiber-wise Calderon-Zygmund decomposition and application to a bi-dimensional paraproduct. Illinois J. Math. 56 (2012), no. 2, 415-422.

[7] J. Bourgain: A remark on the maximal function associated to an analytic vector field. Analysis at Urbana, Vol. I (Urbana, IL, 1986-1987), 111-132.

[8] E. Carneiro and D. Oliveira e Silva: Some sharp restriction inequalities on the sphere. Int. Math. Res. Not. IMRN, to appear.

[9] M, Christ, A. Nagel, E. Stein and S. Wainger: Singular and maximal Radon transforms: analysis and geometry. Ann. of Math. (2) 150 (1999), no. 2, 489-577.

[10] Coifman, Jones and Semmes, Two elementary proofs of the L^2 boundedness of Cauchy integrals on Lipschitz curves. J. Amer. Math. Soc. 2 (1989), no. 3, 553564.

[11] A. Carbery, A. Seeger, S. Wainger and J. Wright: Classes of singular integral operators along variable lines. J. Geom. Anal. 9 (1999), no. 4, 583-605.

[12] C. Demeter and F. Di Plinio: Logarithmic L^p Bounds for Maximal Directional Singular Integrals in the Plane. J. Geom. Anal. 24 (2014), no. 1, 375-416.

[13] D. Foschi: Maximizers for the Strichartz inequality. J. Eur. Math. Soc. (JEMS) 9 (2007), no. 4, 739-774.

[14] S. Guo: Hilbert transform along measurable vector fields constant on Lipschitz curves: L^2 boundedness, arXiv:1401.2890

[15] S. Guo: Hilbert transform along measurable vector fields constant on Lipschitz curves: L^p boundedness, arXiv:1409.3010

[16] P. W. Jones: Square functions, Cauchy integrals, analytic capacity, and harmonic measure. Harmonic analysis and partial differential equations (El Escorial, 1987), 24-68, Lecture Notes in Math., 1384, Springer, Berlin, 1989.

[17] G. Karagulyan: On unboundedness of maximal operators for directional Hilbert transforms. Proc. Amer. Math. Soc. 135 (2007), no. 10, 3133-3141 (electronic).

[18] S. Klainerman and D. Foschi: Bilinear space-time estimates for homogeneous wave equations. Ann. Sci. cole Norm. Sup. (4) 33 (2000), no. 2, 211-274.

[19] V. Kovac: Boundedness of the twisted paraproduct. Rev. Mat. Iberoam. 28 (2012), no. 4, 1143-1164.

[20] M. Lacey and X. Li: Maximal theorems for the directional Hilbert transform on the plane. Trans. Amer. Math. Soc. 358 (2006), no. 9, 4099-4117.

[21] M. Lacey and X. Li: On a conjecture of E. M. Stein on the Hilbert transform on vector fields. Mem. Amer. Math. Soc. 205 (2010), no. 965.

[22] X. Li and C. Muscalu: Generalizations of the Carleson-Hunt theorem. I. The classical singularity case. Amer. J. Math. 129 (2007), no. 4, 983-1018.

[23] M. Lacey and C. Thiele: L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$. Ann. of Math. (2) 146 (1997), no. 3, 693-724.

[24] M. Lacey and C. Thiele: On Caldern's conjecture. Ann. of Math. (2) 149 (1999), no. 2, 475-496.

[25] M. Lacey and C. Thiele: A proof of boundedness of the Carleson operator. Math. Res. Lett. 7 (2000), no. 4, 361-370.

[26] E. M. Stein and B. Street: Multi-parameter singular Radon transforms III: Real analytic surfaces. Adv. Math. 229 (2012), no. 4, 2210-2238.